Bayesian Local Smoothing Modeling and Inference for Pre-surgical FMRI Data

by

Zhuqing Liu

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Biostatistics)
in The University of Michigan
2016

Doctoral Committee:

Professor Timothy D. Johnson, Co-chair
Assistant Professor Veronica J. Berrocal, Co-chair
Professor Michael R. Elliott
Professor Thomas E. Nichols, University of Warwick
Professor Kerby Shedden

To Yi and My Parents

ACKNOWLEDGEMENTS

I would like to gratefully and sincerely thank Dr. Timothy D. Johnson and Dr. Veronica J. Berrocal for their extensive guidance, understanding, and patience during my graduate studies. It is your consistent intellectual and financial support that allowed me to complete my research after a long pursuit. I would also like to acknowledge my committee members: Dr. Michael R. Elliott, Dr. Kerby Shedden and Dr. Thomas E. Nichols for their valuable inputs. In particular, Dr. Andreas Bartsch has been generously sharing his expertise to enrich my research throughout the years of my graduate study.

I would like to give my thanks to all the knowledgeable faculty members in the Department of Biostatistics. From you, I learned devotion, enthusiasm and optimism towards life and research challenges. Especially, I am grateful to Dr. Alexander Tsodikov for being very supportive when I worked as a research assistant under his supervision. I also want to thank all my fellow students and friends in the department and friends in Ann Arbor for all your warm support. All the years being with you will forever stay in my memory.

Finally and most importantly, I would like to thank my parents for their understanding, patience and faith in me. They have always been supportive to my life path. My last and special thank goes to my husband Yi. I want to thank you for all your faithful support in daily life and in research that kept me through my pursuit. Your heartful company during these special and busy years is an essential force driving me forward.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

LIST OF TABLES

Table

ABSTRACT

Bayesian Local Smoothing Modeling and Inference for Pre-surgical FMRI Data

by

Zhuqing Liu

Co-chairs: Dr. Timothy D. Johnson, Dr. Veronica J. Berrocal

There is a growing interest in using fMRI measurements and analyses as tools for pre-surgical planning. One example is to map out essential functional brain regions of specific subjects via fMRI prior to surgery to facilitate tumor resection. For such applications, spatial precision and control over false negatives and false positives are vital, requiring careful design of an image smoothing method and a classification procedure. The standard smoothing method is to convolve the image with a Gaussian kernel, applying a fixed amount of smoothing to the entire image. This approach, however, cannot adapt to the spatially varying signal, causing some regions to be undersmoothed while others oversmoothed with a subsequent loss in spatial precision. In addition, while strict false positive control is usually desired in a neuroscientific fMRI study, controlling false negatives is of greater interest in pre-surgical planning. To this end, this dissertation seeks computationally efficient approaches to overcome the limitation of existing methods and address new challenges in pre-surgical fMRI analyses.

In the first study, we develop a Bayesian solution for the pre-surgical analysis of a single fMRI brain image. Specifically, we propose a novel spatially adaptive con-

ditionally autoregressive model, i.e., the CWAS model, that adaptively and locally smooths the fMRI data. Our model provides more flexibility in characterizing the spatial variation of the fMRI data and introduces a parameter which offers an intuitive interpretation in controlling the degree of smoothing. We introduce a Bayesian theoretical decision approach that allows control of both false positives and false negatives to identify activated and deactivated brain regions. We benchmark the proposed solution to two existing spatially adaptive smoothing models, through simulation studies and two patients' pre-surgical fMRI datasets.

In the second study, we extend the idea of spatially adaptive smoothing to multiple fMRI brain images in order to leverage spatial correlations not only within images but also across multiple images. In particular, we propose three spatially adaptive multivariate conditional autoregressive models that can be considered as extensions of the multivariate conditional autoregressive (MCAR) model (Gelfand and Vounatsou, 2003), the CWAS model, and the model of Reich and Hodges (2008), respectively, and one mixed-effects model assuming that all observed fMRI images originate from one common image. We compare the performance of the proposed models with those from the MCAR and CWAS models using simulation studies and two sets of fMRI brain images, acquired either from the same patient, same paradigm or same patient, different paradigms.

The last study is motivated by fMRI brain images acquired at two different spatial resolutions from the same patient. We develop a Bayesian hierarchical model with spatially varying coefficients to retain the spatial precision from the high resolution image while utilizing information from the low resolution image to improve estimation and inference. Comparisons between the proposed model and the CWAS model, which operates at a single spatial resolution, are performed on simulated data and a patient's multi-resolution pre-surgical fMRI data.

CHAPTER I

Introduction

Since the birth of functional Magnetic Resonance Imaging (fMRI) in 1990, fMRI applications have grown exponentially in both research and industry. Due to its non-invasiveness and wide availability, fMRI has quickly become a popular instrument for researchers to study the brain. Recently, with the wide acceptance and technological improvements of fMRI, it is gaining popularity in clinical settings as well. Pre-surgical planning is one such promising clinical application of fMRI. However, to fully utilize fMRI and thoroughly understand fMRI data, several steps are indispensable, including experimental design, data acquisition, preprocessing, and data analysis. As an interdisciplinary technique, fMRI requires contributions from scientists from various fields. Undoubtedly, statistical considerations and analyses are employed throughout the entire fMRI data analysis. In this dissertation, we focus on developing novel statistical methods to overcome the limitations of existing methods and new challenges brought on by the technological improvements of fMRI in the setting of pre-surgical planning.

The remainder of this chapter is organized as follows: In Section 1.1, we briefly introduce fMRI technology and its use in pre-surgical planning. In Section 1.2, we describe several key steps in fMRI data analysis. We present some existing methods and statistical challenges in Section 1.3 and describe the conditional autoregressive

(CAR) model in Section 1.4. Finally, we outline the dissertation in Section 1.5.

1.1 Overview of fMRI and its use in pre-surgical planning

fMRI is performed by Magnetic Resonance (MR) scanners, a non-invasive medical imaging machine that uses a very strong main magnetic field, gradient fields, and radio frequencies to construct an image of the inside of the body (Revett, 2011). It works by leveraging the fact that certain atoms in the body, e.g. hydrogen, have a net nuclear magnetic moment. When the body is placed in the external magnetic field produced by the MRI scanner, the nuclear spin of the hydrogen atoms tend to align in the direction of the external magnetic field. At room temperature some of the atomic nuclei will align in the opposite direction. There is only about a one-percent difference in number of nuclei that align with the magnetic field as opposed to against it. However, given the sheer number of nuclei, this small difference is large enough to create a net nuclear magnetic moment. A system of orthogonal gradient coils are used to focus the strength of the magnetic field to a very small volume of the body such that only those nuclei in this area are precessing at the Larmor frequency, i.e., the natural resonance frequency of a given nucleus, and are aligned with the external magnetic field. Next, a radio frequency (RF) pulse, orthogonal to the external magnetic field is applied to the system. This causes the nuclei, precessing at the Larmor frequency, to tip in the direction of the RF pulse, orthogonal to the external magnetic field. After the RF pulse is turned off, the nuclear magnetic moments return to precess in the direction of the external magnetic field. A system of radio receiving coils then picks up the electrical signal induced by the change in direction of the magnetic moments, where the strength of the signal received is proportional to the number of nuclei as well as the strength of the external magnetic field. The signal detected lies in the complex plane and the inverse Fourier transform is applied to the signal to create an image of the brain. Please refer to Brown et al. (2014) for a more thorough discussion.

FMRI extends the use of MRI from examining body structure to mapping brain activity. Functional brain activities are detected by changes in Blood Oxygenation Level Dependent (BOLD) signal in response to neuronal activity (Ogawa et al., 1990; Kwong et al., 1992). When a brain region is activated due to some task or external stimulus, blood flow increases in this region changing the ratio of oxygenated to deoxygenated hemoglobin (Springer et al., 1999). Deoxygenated hemoglobin (dHb) is more magnetically susceptible than oxygenated hemoglobin (Hb). The difference in magnetic properties between dHb and Hb leads to local changes in the magnetic fields and ultimately produces the signal which can be captured by the MRI scanner. As the proportion of dHb increases, a loss of magnetization called $T2^{\star}$ decay increases resulting in a loss of signal intensity. This effect increases as the square of the magnetic field (Gomori et al., 1987) and thus relatively high field magnets are preferred (e.g., 3 Tesla vs 1.3 T Tesla).

During a single-subject fMRI study, the subject is asked to alternately perform certain tasks. A typical experiment may alternate between rest periods and "task" periods. During the fMRI data acquisition the entire brain is imaged every 2-4 seconds. This is referred to as the repetition time (TR). Typically, a total of 100-300 images are acquired in a single session. Therefore, a typical fMRI dataset is a multivariate time-series of 3D volumes.

fMRI has been extensively used to study brain function over the last two decades. Although not yet widely accepted for clinical use, fMRI technology is continuously improving and researchers are beginning to explore its potential for clinical applications. One intriguing clinical application of fMRI, and the focus of this dissertation, is its use in pre-surgical mapping of brain functions. A pre-surgical map of brain function could aid neurosurgeons during the planning and resection stages of, say, a tumorectomy (Bartsch et al., 2006; Sunaert, 2006; Bookheimer, 2007; Durnez et al., 2013). By mapping out functional regions of the brain, most commonly motor and

speech functions, pre-surgical fMRI may facilitate the selection of the least harmful treatment and/or surgery that maintains essential brain functions (Stippich, 2007b). Pre-surgical fMRI mapping may also help surgeons in two other ways as well: 1) predict potential deficits in neurofunctions due to surgery or continued lesion growth; 2) optimize the pathway to tumor resection (Håberg et al., 2004; Tieleman et al., 2009). However, pre-surgical planning with fMRI has not yet been adopted by neurosurgeons due to the lack of standardized clinical imaging protocols (Stippich, 2007b).

1.2 Statistical Analysis for fMRI Data

The aim of the statistical analysis for fMRI data is to separate the true signal in response to the stimulus from noise and determine which voxels have associated time-series data patterns that correlate with the known pattern of stimulation. Here we briefly discuss three important steps in fMRI data analyses. See Lindquist (2008) for details.

Pre-processing:

Prior to formal fMRI statistical analyses, fMRI data needs to undergo a series of preprocessing steps in order to remove various types of artifacts introduced during data acquisition and to increase statistical validity for later statistical analyses (Strother, 2006; Lindquist, 2008). Typical preprocessing steps include correction for motion and geometric distortions, image registration and normalization, temporal filtering, spatial smoothing, and pre-whitening. It is common to spatially smooth the fMRI data using a Gaussian kernel with a pre-determined full width at half maximum (FWHM) (Friston et al., 1995). Principally, spatial smoothing is used to increase the signal to noise ratio (SNR). Secondly, it is used to validate the use of Gaussian random field (GRF) theory to control the family-wise error rate (Worsley et al., 1996).

Statistical Analysis: The general linear model (GLM) is the fundamental statistical model used in the analysis of fMRI data (Friston et al., 1994b). It is assumed

that voxels are independent and for each voxel, a GLM is fitted to the voxel's time series data. This method is referred to as the mass univariate approach. Suppose during a single-subject fMRI study, fMRI images consist of N voxels and are collected at T points in time. Let Y_i be the BOLD signal at time i, $i = 1, ..., T$ and let $\boldsymbol{Y}$ be a column vector representing the observed time-series BOLD signal for a voxel, i.e., $\boldsymbol{Y} = (Y_1, Y_2, ..., Y_T)^T$. Suppose there are p predictors. Then the time-series data of each voxel $\boldsymbol{Y}$ is modeled by a GLM which is formulated as:

$$\boldsymbol{Y} \;=\; \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon},$$

where $\boldsymbol{X}$ is the $T \times p$ known design matrix, $\boldsymbol{\beta}$ is the $p \times 1$ vector of unknown parameter coefficients and $\boldsymbol{\epsilon}$ is $T \times 1$ vector of error values.

There are several available methods to estimate the unknown parameter $\boldsymbol{\beta}$. The simplest way is the ordinary least square (OLS) approach which assumes errors are independently and identically distributed, i.e., $\boldsymbol{\epsilon} \sim \mathrm{N}(0, \sigma^2 \boldsymbol{I})$. However, this assumption is violated because at each voxel the data are correlated in time. Bullmore et al. (1996) propose a method which removes the autocorrelation before parameter estimation to obtain the Best Linear Unbiased Estimates (BLUE). This method is also known as the "pre-whitening" method and consists of two steps. In the first step, OLS is used to fit the data and the resultant residuals are used to estimate the autocorrelation which is modeled with some time series model, typically an autoregressive model of order 1 (AR(1)). In the second step, GLM estimation is made on pre-whitened GLM model which is obtained by pre-multiplying both sides of the model by the inverse of the square root of the correlation matrix. Note that spatial correlation among neighboring voxels is ignored in this model fitting step.

After model fitting, hypothesis testing can be performed to test a single contrast of interest using a t-test. The contrast is formulated as $\boldsymbol{c}^T \hat{\boldsymbol{\beta}}$ where $\boldsymbol{c}$ is the contrast

vector. The T-test statistic for $\boldsymbol{c}^T\hat{\boldsymbol{\beta}}$ is given by

$$t = \frac{\boldsymbol{c}^T\hat{\boldsymbol{\beta}}}{\sqrt{\hat{\sigma}^2\boldsymbol{c}^T(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{c}}}.$$

The resulting t-statistic at all voxels forms the statistical parametric map (SPM)—a t-statistic image. Corresponding Z-statistic image can be created by transforming the t-distribution to a normal distribution. Similarly, F-test can be constructed to test multiple contrasts simultaneously. Refer to Penny et al. (2011) for more details.

Adjustment for Multiple Comparisons: Once the SPM is obtained, the next step is to threshold it to determine functionally activated brain regions. This is typically done by classifying the voxels into three classes: activated voxels, deactivated voxels and null voxels. Activated/Deactivated regions or voxels are those regions or voxels where the BOLD signal is significantly larger/smaller during the experimental condition as compared to the control condition. When determining the states for thousands of voxels in one fMRI study simultaneously, multiple comparisons corrections are critical. Standard fMRI studies control for the false discovery rate (Genovese et al., 2002) or the family-wise error rate (Nichols and Hayasaka, 2003). The standard method for controlling the family-wise error rate is based on Gaussian random field theory (Worsley et al., 1996).

1.3 Statistical Challenges in Pre-surgical fMRI Analysis

A large body of research has been published on fMRI data analysis; however, scant literature exists for pre-surgical fMRI data analysis from a statistical point of view. To date, there are no official guidelines for pre-surgical fMRI analysis and its use in medical diagnostics (Stippich, 2007a). When analyzing pre-surgical fMRI data, one should always be aware that the pre-surgical fMRI analysis is fundamentally different from standard fMRI analysis. One major difference is the subject-level sample

size. Standard fMRI analysis typically infers about population level effects incorporating information from a group of healthy or unhealthy subjects (Stippich, 2007b). Pre-surgical fMRI measurements and analyses are always performed on an individual patient (Stippich, 2007a). Here we discuss two main problems we have encountered with pre-surgical data analysis: spatial smoothing and multiple comparisons procedures.

As discussed in Section 1.2, spatial smoothing with a pre-specified Gaussian kernel is an essential step in fMRI data preprocessing. Smoothing with a fixed Gaussian kernel cannot adapt to the spatially varying signal, causing some regions to be undersmoothed while others are oversmoothed (Holmes et al., 1996). Spatial smoothing tends to blur spatial boundaries between activated (or deactivated) regions and null regions. As a result spatial precision between these boundaries is lost. This has potentially detrimental consequences for pre-surgical planning where spatial precision is of utmost importance. Such imprecision may cause some brain regions to be mistakenly disturbed during surgery and thus should be avoided (Yoo et al., 2004).

Standard fMRI analysis is designed to prevent false positives in the hypothesis testing framework. Due to the fact that fMRI data are intrinsically noisy, the true activations are highly likely to be masked by noise. This generates a large number of false negatives. With false positives, certain areas of the brain may be considered as too important to remove, which may result in incomplete removal of brain tumor tissue. With false negatives, essential regions of the brain could be considered as not important and cutting through or removing these areas may lead to irreversible brain damage. Since false negatives may lead to more severe consequences than false positives, pre-surgical fMRI analysis demands a stricter control of false negatives than false positives in the multiple comparison procedure. Thus methods that control false negatives, as well as false positives, should be implemented for pre-surgical fMRI analysis.

To this end, several alternative thresholding methods have been proposed for presurgical fMRI data analysis. Voyvodic et al. (2009) proposed an approach, named activation mapping as a percentage of local excitation (AMPLE), which normalizes t-statistic map by the peak t-statistic in a desired region of interest (ROI) and then thresholds at a predetermined level. Gross and Binder (2014) argued that a ratio of t-values has no obvious meaning or known distribution and the thresholds for the normalized t-statistic map were chosen arbitrarily. Instead, they claimed that the mean response amplitude of the fMRI signal is more relevant for clinical applications than is the t-statistic, since the former is independent of sample size. However, using only the mean amplitude fails to take into account the variability of fMRI data measurements. Further, approach of Gross and Binder (2014) lacks a standard threshold and as a result their threshold values were chosen to match the results from the standard multiple comparison procedures.

1.4 Conditional Autoregressive Model

In this section, we briefly review the conditional autoregressive (CAR) model which will be fully described and extensively used in the remainder of this dissertation.

The conditional autoregressive (CAR) model and the simultaneously autoregressive (SAR) model are the two popular autoregressive models used to model spatial interactions for areal data (Wall, 2004). The SAR and CAR models were originally developed by Whittle (1954) and Besag (1974) respectively. Instead of modeling the covariance structure through the distance between two locations as done for point-level data, both the CAR and SAR models incorporate spatial dependence in the covariance structure via a defined neighborhood system. The most widely used neighborhood structure is the first-order neighborhood system: two areal units are a neighbor-pair if they share a common boundary. The two models can be used to model spatial correlation of the observations directly or they can be introduced

through random effects in hierarchical models (Gelfand and Vounatsou, 2003).

In contrast to the SAR model which is efficient for likelihood methods, the CAR model is computationally convenient for Gibbs sampling and more general Markov Chain Monte Carlo (MCMC) methods (Banerjee et al., 2014). This is because the CAR model, as a special case of a Markov random field, is specified through a set of full conditionals, while Brook's Lemma (Brook, 1964) is invoked to ensure a unique joint distribution. With high dimensional spatial data, it is computationally expensive, if not impossible, to model the covariance matrix directly since matrix inversion or matrix determinant calculation is always involved and the size of the matrix is massive. The CAR model is advantageous compared to the SAR model, or other spatial models, due to its ability to provide immediate posterior full conditional distributions for MCMC sampling. The analysis of imaging data naturally fits into the Bayesian framework where the CAR model can be employed as a prior on the mean structure of the data (Besag, 1993).

1.5 Outline

The rest of the dissertation is organized as follows. Chapter II establishes a Bayesian solution for pre-surgical fMRI analysis including a spatially adaptive conditional autoregressive model that lets the data determine where smoothing should occur and a Bayesian theoretical decision approach to classify the voxels into activated, deactivated and null states. Chapter III proposes three spatially adaptive multivariate conditional autoregressive models and a mixed-effects model to leverage correlation between multiple fMRI images within a single subject. Chapter IV develops a mixed-effects model with spatial varying coefficients model to incorporate multiple fMRI datasets with different resolutions to improve accuracy of the activation/deactivation map. The last chapter concludes our work and provides discussions for future research.

CHAPTER II

Spatially Adaptive Conditionally Autoregressive Models with Application to Univariate FMRI Data

2.1 Introduction

Over the past two decades, researchers in the cognitive neurosciences have used functional magnetic resonance imaging (fMRI) to study changes in brain activation brought on by various experimental stimuli (Ogawa et al., 1990; Kwong et al., 1992). More recently researchers have found fMRI useful in studying functional deficits in patients with neurological diseases or psychiatric disorders, as well as connectivity patterns during rest. In addition, there is growing interest in using fMRI to map out functionally relevant brain regions as a pre-surgical tool to aid neurosurgeons during the planning stages of tumor resection (Bartsch et al., 2006; Sunaert, 2006; Bookheimer, 2007; Durnez et al., 2013). Brain tumor resection is an arduous task. The surgeon's goal is to remove as much tumor as possible while preserving as much healthy brain tissue as possible; especially avoiding the removal, and damage, of healthy regions of the brain that are vital to the patient's quality of life. The neurosurgeon may try to achieve this goal by "awake craniotomy" and electrical stimulation mapping (ESM). That is, once the patient is conscious, the surgeon begins mapping out those regions of the brain near or within the tumor, i.e. peri- or intra-tumor, that

are vital to the patient's quality of life by asking the patient to perform certain tasks. During these tasks, the surgeon probes the brain with an electrode that delivers a small amount of electrical current. Once the task is disrupted, the surgeon places a small marker on that part of the brain that is responsible for the task and will avoid that region during surgery. This surgical mapping and tumor resection may take up to twelve hours or longer depending on the location of the tumor relative to vital brain circuits and regions. These lengthy procedures place great stress on both the patient and the surgeon. Additionally, ESM is subject to both false-negative and false-positive results and is complicated by seizures in up to 15% of the patients (Nossek et al., 2013). Mapping out these vital regions via fMRI prior to surgery allows the surgeon to quickly hone in on these regions for verification, thus speeding up the ESM and tumor resection. Thus, spatial precision of functionally relevant regions, as defined by fMRI scans, is of utmost importance and is the motivation for the work presented in this manuscript.

After acquisition of fMRI data, several steps are typically performed to map brain activity since fMRI data, time series of three-dimensional volume data, each consisting of thousands of uniformly spaced volume elements, called 'voxels' (Lindquist, 2008), are intrinsically noisy. The steps typically are: 1) Pre-processing. This includes correction for motion and geometric distortions, image registration, temporal filtering, spatial smoothing and pre-whitening; 2) Fitting a general linear model to every time series at each voxel, independently of one another. This step is referred to as a mass univariate approach. The resulting image of standardized parameter values is called a Z-statistic image, or map. In reality the resulting standardized parameters all follow a student-T distribution, however, since there are typically hundreds to thousands of observations in each time series at each voxel, the normal distribution is a very good approximation; and 3) Performing hypothesis testing with correction for multiple comparisons. The two most commonly used analysis tools are the Statistical

Parametric Mapping (SPM) (Friston et al., 1994a) software and the FMRIB Software Library (FSL) (Smith et al., 2004; Woolrich et al., 2009). In the pre-processing step, the fMRI data are spatially smoothed by convolving the data with a Gaussian kernel with spatial extent described by the full width of the kernel at half its maximum height (FWHM) (Friston et al., 1995). Spatial smoothing is an important preprocessing step that is performed to achieve two goals. First, it increases the signal to noise ratio. Second, and more importantly, it helps guarantee that the assumptions of Gaussian random field theory, a method commonly used in the following multiple comparisons procedure (Worsley et al., 1996), are satisfied. Although the Gaussian kernel approach is the most popular approach for smoothing fMRI images, Holmes et al. (1996) point out several associated problems, one of which is of great concern for pre-surgical planning: the spatial smoothing pre-processing step uses a Gaussian kernel with a fixed FWHM and does not locally adapt the amount of smoothing based on the data. This can cause some regions to be under-smoothed while others to be over-smoothed, blurring out the boundaries of activated or deactivated regions with null regions. (Null regions or voxels are defined as regions or voxels where the Blood-Oxygenation-Level-Dependent (BOLD) signal (Ogawa et al., 1992) does not change significantly between the fMRI control condition and the fMRI experimental condition. Activated regions or voxels are those regions or voxels where the BOLD signal is significantly larger during the experimental condition as compared to the control condition. Deactivated regions or voxels are those regions or voxels where the BOLD signal is significantly smaller during the experimental condition compared to the control condition. A false negative occurs when a truly activated or deactivated voxel is classified as null while a false positive occurs when a truly null voxel is classified as either activated or deactivated.) Furthermore, for a small (relative to the FWHM of the smoothing kernel) activated or deactivated region, global smoothing can completely obliterate its significance. Thus, one important challenge facing pre-surgical fMRI is to avoid

blurring any sharp boundaries of activated or deactivated regions and to avoid false-negatives by over-smoothing, ensuring spatial accuracy (Yoo et al., 2004).

Our goal in this paper is to develop a statistical model that will adaptively smooth pre-surgical fMRI data; smoothing more in regions where smoothness is warranted and smoothing less in regions where it is not, for example at the interface between activated and null regions. We fit our model to the Z-statistic image in native space. (That is in the geometry of the subject's brain. We do not register to a standard template, however, geometric distortions are corrected, if necessary.) We use the FSL software to perform both the preprocessing step and the mass univariate analysis that results in a Z-statistic image. However, it is important to note that we do not spatially smooth the data in the preprocessing step. A classical statistical method to smooth images is to fit a linear model to the data with spatial random effects using an autonormal model (Besag, 1974). Typically, the spatial random effects are assigned a conditionally autoregressive (CAR) prior (Besag et al., 1991, BYM). In the BYM model, smoothing is controlled by a global smoothing parameter, the spatial variance, and thus a fixed amount of smoothing is applied to the entire image. However, global smoothing is not always warranted and certain parameters in the CAR model can be assigned hyperprior distributions that allow adaptive smoothing. For examples, see Brewer and Nolan (2007, BN) and Reich and Hodges (2008, RH).

Brewer and Nolan (2007) allow for adaptive local smoothing by assigning each site a variance parameter and setting the parameters in the autonormal model in Equation (2.1) to yield Equation (2.7), while Reich and Hodges (2008) set them so to yield Equation (2.8). The two models differ on how the spatially adaptive variances of the full conditionals of the means are defined, see Equations (2.7) and (2.8). Though the RH and BN models are designed to achieve adaptive local smoothing, application of these models to our fMRI data proved problematic. The BN model over-smoothed the Z-statistic image while the RH model failed to converge depending on the initial

state of the Markov chain. These problems are most likely due to the size of the fMRI data (approximately 47000 voxels). We discuss these issues further in Sections 2.4 and 2.5. More recently, Yue et al. (2010) introduced an adaptive spatial smoothing model that uses a non-stationary spatial Gaussian Markov random field and applied it to fMRI data. However, they only consider single slices—two dimensions in space of the 3-dimensional fMRI time series data—applying their model to every slice in the time series (computational considerations prevented them from modeling the full 3-dimensional time series data).

Another approach to avoid blurring of boundaries is to use a Potts model (Potts, 1952). Woolrich et al. (2005) proposed a parametric Potts model for image segmentation that could be used for pre-surgical fMRI. Johnson et al. (2013) proposed a Bayesian non-parametric Potts model and showed that the non-parametric Potts model has better statistical properties if the parametric Potts model is mis-specified, and performed on par with the parametric model when it is correctly specified. While the non-parametric Potts model shows good statistical properties, there are practical issues with computational cost. Johnson et al. (2013) estimate two model parameters in the potential function and both depend on an intractable normalizing constant. Thus, they use a path sampling (Gelman and Meng, 1998) approach to estimate the log ratio of these normalizing constants on a two-dimensional grid of parameter values. The log ratios are determined via simulation in a pre-processing step and this step can take one week or longer. Since pre-surgical fMRI analyses are performed in native space, this pre-processing step must be performed for each individual. A possible delay in surgery to obtain the final analysis is not acceptable. Recently, we learned of a method (Murray et al., 2006) that can estimate these parameters in a Potts model without the pre-processing step. However, we did not find any advantage with respect to computation time over the method proposed by Johnson et al. (2013). Therefore, we propose a novel model that adaptively and locally smooths the

fMRI data. Compared to the BN and RH models, our model offers a more intuitive interpretation and presents less problems during model fitting. In particular, in our model we let the variances in the full conditional of the means be proportional to the error variances, allowing the degree of smoothing to vary across the brain. We call our model a conditionally weighted adaptive smoothing model (CWAS, see Section 2.2.2.1 for details).

In pre-surgical fMRI data, false negatives (i.e., voxels falsely classified as null) are more dangerous to the patient than false positives (i.e., voxels falsely classified as activated or de-activated) since the former may result in the surgeon damaging healthy tissue vital to quality of life (Haller and Bartsch, 2009). Building upon the work of Müller et al. (2007), we propose a loss function to identify functionally active regions in the brain that asymmetrically penalizes false negatives and false positives. We use the proposed loss function to compare the performance of our model to that of the BN and RH models. Simulation studies show that the proposed model outperforms the BN and RH models in terms of false negative rates.

The remainder of this paper is organized as follows. We begin in Section 2.2 with a description of the proposed CWAS model, a brief overview of the BN and RH models, and our proposed loss function. We then discuss model implementation details in Section 2.3. In Section 2.4, we present results from our simulation studies. We then apply our model to two pre-surgical fMRI data sets in Section 2.5 and compare results with those from the BN and RH models. We conclude the paper in Section 2.6 with a discussion of the contribution of this paper and future directions.

2.2 Model

2.2.1 The CAR model

Consider an unsmoothed 3-dimensional Z-statistic image. We use $\boldsymbol{Y}$ to represent its Z-statistic intensity for the set of N voxels, where Y_i is the intensity of the ith voxel, for $i = 1, ..., N$. We assume a first order neighborhood system with two voxels, i and i', neighbors if they share a common face and we denote it by $i \sim i'$. Suppose μ_i and ϵ_i are the mean intensity and the random measurement error of voxel i, respectively, with error $\epsilon_i \sim \mathrm{N}(0, \sigma_i^2)$, where $\mathrm{N}(a, b)$ denotes the normal distribution with mean a and variance b. Denote the mean intensity vector by $\boldsymbol{\mu} = (\mu_1, ..., \mu_N)^T$. In this context, $\boldsymbol{Y}$ is the Z-statistic image and $\boldsymbol{\mu}$ represents the smoothed Z-statistic image. We assume $Y_i = \mu_i + \epsilon_i$, i.e., $Y_i \sim \mathrm{N}(\mu_i, \sigma_i^2)$.

A classical statistical model to smooth images is the CAR model (Besag et al., 1991). The CAR model is a specific form of the auto-normal model (Besag, 1974), the latter being defined by the set of full conditionals:

$$[\mu_i \mid \boldsymbol{\mu}_{(-i)}, \boldsymbol{B}, \phi_i^2] \sim \mathrm{N}\left(\sum\nolimits_j b_{ij}\mu_j,\ \phi_i^2\right) \tag{2.1}$$

where $\boldsymbol{\mu}_{(-i)}$ is the $n-1$ dimensional vector obtained from $\boldsymbol{\mu}$ by removing μ_i, and $\boldsymbol{B}$ is an $N \times N$ matrix with elements b_{ij}, where $b_{ij} = 0$ if $i = j$.

Set $w_{ij} = 1$ if and only if voxels i and j are neighbors (note: a voxel is not a neighbor of itself), set $w_{i+} = \sum_{j=1}^{N} w_{ij}$, and let $\boldsymbol{W}$ be the $N \times N$ matrix with elements w_{ij}. The CAR model is obtained when $b_{ij} = w_{ij}/w_{i+}$ and $\phi_i^2 = \tau^2/w_{i+}$ with $\tau \in \mathbb{R}^+$. Then, the full conditionals in Equation (2.1) are:

$$[\mu_i \mid \boldsymbol{\mu}_{(-i)}, \boldsymbol{W}, \tau^2] \sim \mathrm{N}\left(\sum\nolimits_j (w_{ij}/w_{i+})\,\mu_j,\ \tau^2/w_{i+}\right). \tag{2.2}$$

Here $\sum_j (w_{ij}/w_{i+})\,\mu_j$ indicates the average of the μ_js at voxels that are neighbors

to voxel i and thus the resulting μ_i is smoothed towards the mean of its neighbors. The amount of smoothing in Equation (2.2) is controlled by a global parameter τ^2 while in the original auto-normal model the amount of smoothing is controlled by the voxel-specific variance ϕ_i^2.

By Brook's Lemma (Brook, 1964), the full conditionals in Equation (2.2) imply the following joint prior on $\boldsymbol{\mu}$ up to a normalizing constant:

$$\pi(\boldsymbol{\mu} \mid \tau^2) \propto \exp\left\{-0.5\tau^{-2} \sum\nolimits_{i\sim j} (\mu_i - \mu_j)^2\right\},$$

which is also called pairwise difference prior since it only depends on the differences of all neighbor pairs (Besag, 1993).

2.2.2 Spatially adaptive CAR models

2.2.2.1 The CWAS model

We define our spatially adaptive CAR model on $\boldsymbol{\mu}$ by specifying a specific form for the variance ϕ_i^2 in Equation (2.1). Instead of using a global parameter τ^2 for all voxels as in the CAR model, we let ϕ_i^2 vary across the brain and model it to be proportional to the error variance σ_i^2, i.e., $\phi_i^2 = c_i\sigma_i^2$, where $c_i > 0$. We still assume $b_{ij} = w_{ij}/w_{i+}$. Then, our spatially adaptive CAR model is defined through the following conditional distributions:

$$[\mu_i \mid \boldsymbol{\mu}_{(-i)}, \boldsymbol{W}, c_i, \sigma_i^2] \sim \mathrm{N}\left(\sum\nolimits_j (w_{ij}/w_{i+})\, \mu_j,\, c_i\sigma_i^2\right),\ i = 1, \ldots, N. \tag{2.3}$$

To allow for more flexibility, we let the error variances vary spatially as well by placing a CAR prior on their logarithm, that is:

$$[\ln(\sigma_i^2) \mid \ln(\boldsymbol{\sigma}_{(-i)}^2), \boldsymbol{W}, \lambda^2] \sim \mathrm{N}\left(\sum\nolimits_j (w_{ij}/w_{i+}) \ln(\sigma_j^2),\, \lambda^2/w_{i+}\right). \tag{2.4}$$

Equations (2.6) and (2.4) are the motivation for calling our model the conditionally weighted adaptive smoothing (CWAS) model.

The entire CWAS model is as follows:

$$\begin{aligned}
[y_i \mid \mu_i, \sigma_i^2] &\overset{ind}{\sim} \mathrm{N}\left(\mu_i,\, \sigma_i^2\right),\ i = 1, ..., N \\
\left[\mu_i \mid \boldsymbol{\mu}_{(-i)}, c_i, \sigma_i^2\right] &\sim \mathrm{N}\left(\sum\nolimits_j (w_{ij}/w_{i+})\, \mu_j,\, c_i\sigma_i^2\right), i = 1, ..., N \qquad (2.5) \\
\left[\ln(\sigma_i^2) \mid \ln(\boldsymbol{\sigma}_{(-i)}^2), \boldsymbol{W}, \lambda^2\right] &\sim \mathrm{N}\left(\sum\nolimits_j (w_{ij}/w_{i+})\, \ln(\sigma_j^2),\, \lambda^2/w_{i+}\right), i = 1, ..., N \\
\lambda^2 &\sim \mathrm{IG}\,(a,\, b) \\
c_i &= p_i/(1-p_i), i = 1, ..., N \\
p_i &\overset{iid}{\sim} \mathrm{Beta}\,(\alpha,\, \beta)\,,
\end{aligned}$$

with hyperprior parameters $a = b = 1, \alpha = \beta = 2$. $\mathrm{IG}(\cdot, \cdot)$ denotes the inverse gamma distribution and $\mathrm{Beta}(\cdot, \cdot)$ denotes the beta distribution. To simplify notation, we will denote $\tilde{\mu}_i = \sum_j (w_{ij}/w_{i+})\, \mu_j$. Let $p_i = c_i/(1 + c_i)$, then the full conditional for μ_i is

$$\left[\mu_i \mid y_i, \boldsymbol{\mu}_{(-i)}, \sigma_i^2, p_i\right] \sim \mathrm{N}\left(p_i y_i + (1 - p_i)\tilde{\mu}_i,\, p_i\sigma_i^2\right). \qquad (2.6)$$

Each p_i has an intuitive interpretation in our context: it is the parameter that controls the amount of smoothing at voxel i. If $p_i > 0.5$, more weight is placed on y_i, resulting in relatively less smoothing, while if $p_i < 0.5$ more weight is placed on $\tilde{\mu}_i$, resulting in relatively more smoothing. We will discuss our hyperprior choices in Section 2.4.

2.2.2.2 The BN and RH models

We now explain the differences between our proposed CWAS model and the two spatially adaptive models proposed by Brewer and Nolan (2007, BH) and Reich and Hodges (2008, RH), and our motivation for proposing a new model.

In the BN model, instead of using spatially varying error variances, the authors

assume that $y_i = \mu_i + \epsilon_i$, with $\epsilon_i \sim \mathrm{N}(0, \sigma^2)$. In addition, in the BN model the prior placed on μ_i is given by Equation (2.1) with $b_{ij} = w_{ij} \left\{ (\tau_i^2 + \tau_j^2) \left[\sum_{i \sim j} (\tau_i^2 + \tau_j^2)^{-1} \right] \right\}^{-1}$ and $\phi_i^2 = \left[\sum_{i \sim j} \left(\tau_i^2 + \tau_j^2 \right)^{-1} \right]^{-1}$. Specifically, the conditional distribution for μ_i in Equation (2.1) is given by:

$$\left[\mu_i \mid \boldsymbol{\mu}_{(-i)}, \boldsymbol{\tau}^2 \right] \sim \mathrm{N} \left(\sum\nolimits_{i \sim j} \frac{\mu_j}{\tau_i^2 + \tau_j^2} \left(\sum\nolimits_{i \sim j} \frac{1}{\tau_i^2 + \tau_j^2} \right)^{-1}, \left(\sum\nolimits_{i \sim j} \frac{1}{\tau_i^2 + \tau_j^2} \right)^{-1} \right), \tag{2.7}$$

where $\boldsymbol{\tau}^2 = (\tau_1, \ldots, \tau_N)$.

In the BN model, the authors use an empirical Bayesian method to specify the hyperprior distributions on the τ_i^2s. They first fit the BYM model to their data to get raw estimates of the μ_is. Then, they set the mean of the prior for each τ_i^2 to be

$$\begin{aligned} \theta_i &= w_{i+} \mathrm{Var}(\mu_i - \mu_j)/2 \\ &= w_{i+} \left[\sum\nolimits_{j \in N(i)} (\mu_i - \mu_j)^2 / w_{i+} - (\sum\nolimits_{j \in N(i)} (\mu_i - \mu_j)/w_{i+})^2 \right] /2 \,, \end{aligned}$$

and implement their spatially adaptive smoothing model to get the final estimate of μ_i. When we apply the BN model in our simulation study and to our data, we set $\tau_i^2 \sim \mathrm{IG}(2, \theta_i)$, for each i.

Reich and Hodges (2008) take a different approach by setting b_{ij} and ϕ_i^2 in Equation (2.1) to $w_{ij} \tau_j^{-1} / \sum_{i \sim k} \tau_k^{-1}$ and $\tau_i / \sum_{i \sim j} \tau_j^{-1}$, respectively. In the RH model the conditional distributions of μ_i, $i = 1, \ldots, N$ are

$$\left[\mu_i \mid \boldsymbol{\mu}_{(-i)}, \boldsymbol{\tau}^2 \right] \sim \mathrm{N} \left(\sum\nolimits_{i \sim j} \left(\tau_j^{-1} / \sum\nolimits_{i \sim k} \tau_k^{-1} \right) \mu_j, \, \tau_i / \sum\nolimits_{i \sim j} \tau_j^{-1} \right). \tag{2.8}$$

Similar to our CWAS model, in the RH model the authors assume that $y_i = \mu_i + \epsilon_i$, with $\epsilon_i \sim \mathrm{N}(0, \sigma_i^2)$. Both the σ_i^2s and τ_i^2s vary spatially with CAR priors placed on

their logarithms:

$$\begin{aligned}\left[\ln(\sigma_i^2) \mid \ln(\boldsymbol{\sigma}^2_{(-i)}), \boldsymbol{W}, \lambda^2\right] &\sim \mathrm{N}\left(\sum\nolimits_j (w_{ij}/w_{i+}) \ln(\sigma_j^2),\ \lambda^2/w_{i+}\right), \\ \left[\ln(\tau_i^2) \mid \ln(\boldsymbol{\tau}^2_{(-i)}), \boldsymbol{W}, \gamma^2\right] &\sim \mathrm{N}\left(\sum\nolimits_j (w_{ij}/w_{i+}) \ln(\sigma_j^2),\ \gamma^2/w_{i+}\right), \end{aligned} \tag{2.9}$$

where λ^2 and γ^2 are both assumed to follow an inverse Gamma distribution: IG(1,1).

Our model is different from both the BN model and the RH model in the definitions of the b_{ij} and ϕ_i^2 in Equation (2.1). We believe that compared to the BN and RH models our model offers a more intuitive interpretation. The model parameter p_i for voxel i is the weight placed on the data y_i and controls the amount of smoothing in the CWAS model at voxel i.

2.2.3 Loss function

The goal of using fMRI for pre-surgical planning is to identify which voxels did not activate or deactivate and are safe to resect. After fitting our CWAS model to data, we use a Bayesian decision theoretic approach to classify voxels into two classes, null and non-null (those voxels that are either activated or deactivated). A typical fMRI data analysis places emphasis on only controlling the type I error. However, in pre-surgical fMRI planning, surgeons are more concerned with false negatives since damage to false negative regions of the brain may lead to irreversible deficits in function. Therefore, we propose a novel loss function that explicitly controls both false negatives and false positives, while at the same time allowing false negatives and false positives to be treated asymmetrically in the loss function. Our loss function is a generalization of a loss function proposed by Müller et al. (2007). Let μ_i indicate the mean intensity for voxel i, $r_i \in \{0, 1\}$ denote the true binary state of voxel i (0 for null, 1 for non-null), and $\delta_i \in \{0, 1\}$ represent the estimated state. Let $\boldsymbol{\mu} = (\mu_1, ..., \mu_N)$ and $\boldsymbol{\delta} = (\delta_1, ..., \delta_N)$ be the vectors corresponding to all voxels and let k_1, k_2 and t be

positive weights. Then our proposed loss function is defined as follows:

$$\mathrm{L}(\boldsymbol{\mu}, \boldsymbol{\delta}) = \sum\nolimits_i \left[-r_i\delta_i - (1-r_i)(1-\delta_i) + k_1 r_i(1-\delta_i) + k_2(1-r_i)\delta_i + t\delta_i\right].$$

The terms in this loss function indicate: i) two types of gains corresponding to correct classification of an activated or deactivated voxel, $-r_i\delta_i$, and of a null voxel, $-(1-r_i)(1-\delta_i)$; ii) two types of losses corresponding to incorrect classification of an activated or deactivated voxel, $k_1 r_i(1-\delta_i)$, and of a null voxel, $k_2(1-r_i)\delta_i$; and iii) a penalty term for the total number of discoveries ensuring parsimony, $\sum_i t\delta_i$. Above, k_1 is a weight for incorrectly classified non-null voxels (false negatives) while k_2 is a weight for incorrectly classified null voxels (false positives). When $k_1 > k_2$ more weight is given to incorrectly classified activated voxels relative to null voxels.

Given the data, the optimal decision rule is obtained when the posterior expected loss

$$\begin{aligned}\mathbb{E}\left(\mathrm{L}(\boldsymbol{\mu}, \boldsymbol{\delta} \mid \boldsymbol{Y})\right) &= \sum\nolimits_i \left[-(2+k_1+k_2)\delta_i \Pr(r_i = 1 \mid \boldsymbol{Y}) + (1+k_1)\Pr(r_i = 1 \mid \boldsymbol{Y})\right. \\ &\quad \left. + (1+k_2+t)\delta_i - 1\right]\end{aligned}$$

is minimized. The optimal decision is

$$\delta_i^{\dagger} = \mathrm{I}\left[\Pr(r_i = 1 \mid \boldsymbol{Y}) \geq (1+k_2+t)/(2+k_1+k_2)\right].$$

However, we do not explicitly define $\Pr(r_i = 1 \mid \boldsymbol{Y})$ in our model. Instead, we let $m_i = \left|\mathbb{E}(\mu_i \mid \boldsymbol{Y})/\sqrt{\mathrm{Var}(\mu_i \mid \boldsymbol{Y})}\right|$ and consider for each voxel i

$$f(m_i) = f\left(\left|\mathbb{E}(\mu_i \mid \boldsymbol{Y})/\sqrt{\mathrm{Var}(\mu_i \mid \boldsymbol{Y})}\right|\right),$$

where f is some monotone function and m_i can be considered as the strength of a

voxel being non-null (i.e. either activated or deactivated); with $m_i = 0$ if $r_i = 0$ and $m_i > 0$ when $r_i = 1$ (see Müller et al. (2007) for details). Our loss function is now

$$\begin{aligned} \mathrm{L}(\mathbf{m}, \boldsymbol{\delta}) &= \sum\nolimits_i \{-f(m_i)\delta_i - [1 - f(m_i)](1 - \delta_i) + k_1 f(m_i)(1 - \delta_i) \\ &\quad + k_2[1 - f(m_i)]\delta_i + t\delta_i\} \end{aligned}$$

where $\mathbf{m} = (m_1, m_2, \ldots, m_N)$ and the optimal decision is

$$\delta_i^\star = \mathrm{I}\left[f(\widehat{m}_i) \geq (1 + k_2 + t)/(2 + k_1 + k_2)\right]$$

and $\widehat{m}_i$ is the estimated posterior of m_i. For the function f, a reasonable choice is that by Gross and Binder (2014): that is, $f(m_i) = m_i/q_\alpha(\mathbf{m})$ where q_α is the $(1 - \alpha)$ quantile function of its argument. By choosing an appropriate level of α one can guard against outliers. In our application, there does not appear to be any outliers and so we choose $\alpha = 0$. That is, $q_0(\mathbf{m}) \equiv \max(\mathbf{m})$. Another appropriate choice would be $\alpha = 0.01$. The decision rule of our proposed loss function only depends on the values of the constants k_1, k_2 and t. Once fixed, we have a uniform decision rule that we can apply across different data sets.

2.3 Posterior estimation

2.3.1 The CWAS model

We now discuss posterior estimation for the parameters in our model. We implement a hybrid Metropolis-within-Gibbs algorithm to sample the parameters from their full conditionals. The full conditionals for $\mu_i,\ i = 1, ..., N$ and λ^2 are available in closed form and are respectively, the normal distribution given in Equation (2.6),

and an inverse gamma distribution

$$\left[\lambda^2 \mid \boldsymbol{\sigma}^2\right] \sim \text{IG}\left(a + 0.5(N-1),\, b + 0.5\sum\nolimits_{i\sim j}\left[\ln(\sigma_i^2) - \ln(\sigma_j^2)\right]^2\right).$$

The prior and posterior for both σ_i^2 and c_i are not conjugate pairs. The full conditionals, up to a constant of proportionality, are:

$$\pi(\sigma_i^2 \mid y_i, \boldsymbol{\sigma}^2_{(-i)}, \boldsymbol{\mu}, \lambda^2) \;\propto\; \pi(y_i \mid \mu_i, \sigma_i^2)\pi(\sigma_i^2 \mid \boldsymbol{\sigma}^2_{(-i)}, \lambda^2)\pi(\boldsymbol{\mu} \mid \boldsymbol{c}, \boldsymbol{\sigma}^2), \tag{2.10}$$

$$\pi(c_i \mid \boldsymbol{c}_{(-i)}, \boldsymbol{\mu}, \boldsymbol{\sigma}^2) \;\propto\; \pi(c_i)\pi(\boldsymbol{\mu} \mid \boldsymbol{c}, \boldsymbol{\sigma}^2), \tag{2.11}$$

where the prior distribution of c_i is induced by the prior distribution on p_i. The prior density of c_i is

$$\pi(c_i) = c_i^{\alpha-1}(1+c_i)^{-(\alpha+\beta)}/\text{B}(\alpha,\beta),$$

where $\text{B}(\cdot)$ denotes the Beta function. We remark that although the c_is are a priori independent, they are spatially correlated in the posterior. The full conditional of c_i is

$$\pi(c_i \mid \boldsymbol{c}_{(-i)}, \boldsymbol{\mu}, \boldsymbol{\sigma}^2) \propto c_i^{-\frac{1}{2}} \exp\left\{-\left(\mu_i - \tilde{\mu}_i\right)^2/(2c_i\sigma_i^2)\right\} c_i^{\alpha-1}(1+c_i)^{-(\alpha+\beta)}.$$

Note that this full conditional depends on $\tilde{\mu}_i = \sum_j (w_{ij}/w_{i+})\mu_j$—thus borrowing information from neighboring voxels. Furthermore, the posterior of c_i depends on μ_i and σ_i^2 which are explicitly dependent on their neighbors (both a priori and a posteriori). Thus, the c_is (and p_is) are a posteriori correlated and borrow strength implicitly from their neighboring values. In Section 2.6, we illustrate how one can modify the priors on the p_is so that they are spatially dependent a priori, which may be desirable.

The full conditionals in Equations (2.10) and (2.11) are intractable due to the fact that $\pi(\boldsymbol{\mu} \mid \boldsymbol{c}, \boldsymbol{\sigma}^2)$ is intractable. By Brook's Lemma (Brook, 1964), the prior of $\boldsymbol{\mu}$ is guaranteed to exist. However, with our specification of the conditional priors on

the μ_is in Equation (2.5), the prior of $\boldsymbol{\mu}$ does not have a tractable density. To overcome this issue, we use the pseudo-likelihood approach (Besag, 1975) to approximate the prior of $\boldsymbol{\mu}$, $\pi(\boldsymbol{\mu} \mid \boldsymbol{c}, \boldsymbol{\sigma}^2)$. The pseudo-likelihood approximation was originally developed for spatial models by Besag (1975) motivated by the computational intractability of the true likelihood. Besag (1977) showed that parameter estimation using the pseudo-likelihood approximation is efficient for simple Gaussian fields. The pseudo-likelihood approximation proposed by Besag (1975) for $\boldsymbol{\mu}$ is formulated as the product of all the full conditionals $\pi(\mu_i \mid \cdot)$:

$$\pi(\boldsymbol{\mu} \mid \boldsymbol{c}, \boldsymbol{\sigma}^2) \approx \prod_{i=1}^{N} (2\pi c_i \sigma_i^2)^{-\frac{1}{2}} \exp\left\{ -\left(\mu_i - \tilde{\mu}_i\right)^2 / (2 c_i \sigma_i^2) \right\},$$

where $\boldsymbol{c} = (c_1, ..., c_N)^T$ and $\boldsymbol{\sigma}^2 = (\sigma_1^2, ..., \sigma_N^2)^T$. Since the full conditionals of σ_i^2 and c_i, $i = 1, ...N$ do not have closed forms, we draw samples from their full conditionals using the Metropolis-Hastings algorithm (Hastings, 1970). Note that only the estimation of the c_is and the σ_i^2s rely on the pseudo-likelihood approximation.

2.3.2 The BN and RH models

Brewer and Nolan (2007) also used the pseudo-likelihood approximation to the prior of $\boldsymbol{\mu}$ in their model. In their case, the prior for $\boldsymbol{\mu}$ is approximated as

$$\pi(\boldsymbol{\mu} \mid \boldsymbol{\tau^2}) \approx \prod_{i=1}^{N} \left(\frac{2\pi}{\sum_{i\sim j} \left(\tau_i^2 + \tau_j^2\right)^{-1}} \right)^{-\frac{1}{2}} \exp\left\{ - \frac{\left(\mu_i - \sum_{i\sim j} \frac{\mu_j}{\tau_i^2+\tau_j^2} \Big/ \sum_{i\sim j} \frac{1}{\tau_i^2+\tau_j^2} \right)^2}{2\left[\sum_{i\sim j} \left(\tau_i^2 + \tau_j^2\right)^{-1}\right]^{-1}} \right\}.$$

Reich and Hodges (2008) use the following joint prior distribution for $\boldsymbol{\mu}$:

$$\pi(\boldsymbol{\mu} \mid \boldsymbol{\tau}^2) \propto |\boldsymbol{Q}|^{\frac{1}{2}} \exp\left\{ -0.5 \sum\nolimits_{i\sim j} \left(\mu_i - \mu_j\right)^2 / (\tau_i \tau_j) \right\},$$

where the matrix $\boldsymbol{Q}$ has entries $Q_{ii} = \sum_{i\sim j} \tau_{ij}^{-2}$ and $Q_{ij} = -I(i \sim j)\tau_{ij}^{-2}$, while the

determinant of $\boldsymbol{Q}$ is calculated as the product of all positive eigenvalues of the non-negative definite matrix $\boldsymbol{Q}$. For our Z-statistic image with approximately 47000 voxels (after skull stripping and exclusion of non-brain voxels) computing the determinant of $\boldsymbol{Q}$ is computationally infeasible; therefore, when we apply the RH model to our Z-statistic image, we implement a pseudo-likelihood approach to approximate the prior of $\boldsymbol{\mu}$ with:

$$\pi(\boldsymbol{\mu} \mid \boldsymbol{\tau^2}) \approx \prod_{i=1}^{N} \left(\frac{2\pi\tau_i}{\sum_{i\sim j} \tau_j^{-1}} \right)^{-\frac{1}{2}} \exp\left\{ -\left(\mu_i - \sum\nolimits_{i\sim j} \frac{\tau_j^{-1}\mu_j}{\sum_{i\sim k} \tau_k^{-1}} \right)^2 / \left(\frac{2\tau_i}{\sum_{i\sim j} \tau_j^{-1}} \right) \right\}.$$

2.4 Simulation studies

We conducted simulation studies to investigate the performance of the proposed CWAS model. In the simulation studies, we i) investigate sensitivity to different prior specifications on the p_is; ii) demonstrate the relationship between the number of correctly classified activated voxels and the thresholds used to detect activation regions for the simulated fMRI data with different signal to noise ratios; and iii) compare the performance of our proposed CWAS model with that of the BN and RH models.

For simulation purposes, we used the neuRosim package described in Welvaert et al. (2011) to simulate a set of four-dimensional fMRI data. The detailed settings used for data generation are elaborated as follows. A box-car block design was used for the stimulus function, with each on/off episode lasting 20 seconds. Repetition time (TR) is set to two seconds with a total of 100 scans in each simulated data set. Thus each simulated experiment lasts 200 seconds, with activation onset at 21,61,101,141,181 seconds. The stimulus function is convolved with a gamma hemodynamic response function (HRF). The brain mask for the simulated data is chosen to be the same as that of the patient described in Section 2.5. The brain mask is a three-dimensional $64 \times 64 \times 40$ binary array indicating whether an element in the array

corresponds to a voxel in the brain. There are six spherical activation regions with different sizes and HRF signal magnitudes. Details of the six activation regions are shown in Table 2.1. The baseline signal is set to 250 and the signal fading rate is set to 0.01 as suggested by Welvaert et al. (2011). The noise is a mixture of Gaussian white noise (30%), and spatial noise (70%) with default auto-correlation coefficient equal to 0.75. We simulated 50 data sets for each signal to noise ratio (SNR) considered (SNR = 1, 2 and 3). The average SNR is defined as SNR= $\overline{S}/\sigma_N$, where $\overline{S}$ represents the average magnitude of the signal, and σ_N stands for the standard deviation of the noise (Krüger and Glover, 2001). Z-statistic images were created using FEAT within the FSL (Smith et al., 2004) software package with spatial smoothing turned off in the preprocessing stage. Based on these settings, each Z-statistic image contains a total of 46932 voxels, 328 of which are truly active.

The 150 simulated data sets were then fitted with the CWAS, RH, and BN models. For all three models, we ran the MCMC algorithm for 150,000 iterations with the first 100,000 iterations discarded as burn-in. The RH model encountered convergence problems that might be attributable to i) the large data size, or ii) the pseudo-likelihood approximation. However, when the RH model did converge, our model always converged faster. We visually assessed the traceplots based on which we concluded our model converged faster. Convergence was further verified using the Gelman-Rubin convergence diagnostic statistic (Gelman and Rubin, 1992; Brooks and Gelman, 1998). For cases when the RH model did not converge, we reran the MCMC simulations using a different starting value so that all models were compared using the same 150 simulated data sets.

To investigate sensitivity to different prior specifications under the CWAS model, we considered three Beta distributions, Beta(2,2), Beta(1,3) and Beta(3,1), as the hyperprior distribution of the p_is. We also attempted to control the amount of smoothing in the RH model by implementing different hyperpriors on λ^2 or γ^2 (see Equation

(2.9)). However, smoothing was rather insensitive to the prior information unless we specified a very tight prior. A similar phenomenon was observed for the BN model. We applied our proposed loss function and summarized the false negative rate (FNR) for all three models. Instead of using a fixed threshold derived by minimizing the posterior expected loss and holding the constants k_1, k_2 and t in the loss function fixed, we varied the threshold on $f(m_i)$ from 0 to 1 to demonstrate the trade-off between threshold and true positive counts. For each SNR, Figure 2.1 shows the average FNR over the 50 simulated data sets versus threshold. Each curve denotes a different amount of smoothing for the CWAS model. Beta(2,2), Beta(3,1) and Beta (1,3) are different priors for the smoothing parameters p_i, which represents the weight assigned to the data point y_i: a larger p_i implies less smoothing at voxel i. As shown in Figure 2.1, the curve corresponding to Beta(1,3) results in the largest amount of smoothing and has the lowest FNR compared to the other two priors at a SNR of 1. On the other hand, the curve corresponding to Beta(1,3) has the largest FNR when the SNR is 3. This is consistent with our intuition that less smoothing is necessary as the SNR increases. We also note that our model has a consistently lower FNR than the other two models for a large range of thresholds.

We compared the smoothed Z-statistic images derived from the three models applied on one selected simulated data set. Figure 2.2 shows the comparison between the posterior estimates of the mean intensities from the three spatially adaptive smoothing models when SNR=2. Similar results, not shown, were obtained for signal-to-noise ratios equal to 1 and 3 as well. The top row of Figure 2.2 displays the true activation regions in four sagittal slices of the brain, while the second row presents the simulated data with noise added. By comparing row 3 through 5, it is evident that the amount of smoothing varies among the three models. The BN model severely smoothes the data. The RH model also tends to oversmooth the data. Our model smoothes the simulated data the least and it achieves the goal of balancing smoothing and retaining

relatively sharp boundaries between activated and non-activated areas.

2.5 Application to pre-surgical fMRI data

The motivation of this work is to establish and test the performance of a novel, spatially adaptive smoothing model for pre-surgical fMRI data. In this section we illustrate the application of our proposed CWAS model to the fMRI data from two brain tumor patients and compare it with the RH and BN models. Both patients were pre-surgically mapped by fMRI to determine the peri- and potentially intra-tumoral regions of the brain vital for speech and language functions in order to optimize a) access to the tumor, b) intra-operative ESM at a limited number of predefined cortical sites and c) extent of resection without inflicting new aphasic deficits.

The first patient's tumor was located in proximity to the so-called dorsal stream of speech and language while the second patient's tumor was in proximity to the so-called ventral stream (Hickok and Poeppel, 2007). Consequently, two different fMRI paradigms were applied, specifically challenging functions of the dorsal and ventral stream (see below). A 30 seconds ON/OFF-boxcar block design was used in both cases, and task speed was adjusted to the individual's optimal performance level. Data on the first patient are from a thirty-two year old, right-handed woman who had initially experienced a seizure and was then brought into the emergency room, still with difficulties speaking. In particular, she had persistent problems repeating phonemically challenging phrases. In the fMRI paradigm the patient was therefore instructed and cued to alternate between the silent recitation of challenging tongue twister phrases (ON) and covert repetition of the unchallenging, rhythmic phoneme sequence "tock-tock-tock" (OFF). After partial resection, the tumor, in the left insular and inferior frontal lobe, was classified as an oligodendroglioma with anaplastic components (see the structural MRI scan in the left panel of Figure 2.3 and Figure 2.4). Data on the second patient are from a sixty-two year old, right-handed woman

who presented difficulties with word finding, comprehension and reading but who did not display signs of agraphia. In the fMRI paradigm the patient was therefore asked to silently read non-final embedded clause sentences (ON) alternating with consonant strings (OFF). After partial resection, the tumor, located in the left middle and inferior temporal gyrus, was classified as a glioblastoma multiforme (GBM) with intra-tumoral hemorrhages (see Figure 2.5).

Both fMRI studies were recorded on a 3 Tesla TimTrio scanner (TQ engine, 32 channel head coil; Siemens Medical Solutions, Erlangen) using gradient-echo (GE) echo-planar imaging (EPI) at the following parameters: TR = 3000ms, TE = 30ms, echo spacing = 0.69ms, GRAPPA acceleration factor 2, $n = 160$ volumes in time, $64 \times 64 \times 40$ voxels in space ($3.00 \times 3.00 \times 3.00\text{mm}^3$ resolution, with an additional 15% inter-slice gap). Geometric distortions were corrected using an additional GE fieldmap scan. High-resolution anatomical T2-weighted FLAIR- (fluid-attenuated inversion recovery) and T1-weighted MPRAGE (magnetization-prepared rapid gradient echo) sequences were recorded as well for intra-operative neuro-navigation, cortical surface reconstruction and the display of fMRI results (cf. Figures 2.4 and 2.5). A distortion-corrected 3-dimensional Z-statistic image was estimated in native GE-EPI space ($64 \times 64 \times 40$ voxels) for each patient using the FEAT analysis tool v.6.00 (with convolution of the design matrix by the canonical double-gamma hemodynamic response function, HRF, and inclusion of temporal derivatives) in FSL 5.0.7 (http://www.fmrib.ox.ac.uk/fsl/, Smith et al. (2004); Woolrich et al. (2009)). Spatial smoothing was turned off during the preprocessing stage, motion correction was performed and data were temporally filtered using a high-pass cutoff of 90 seconds (default Gaussian-weighted least-squares straight line fitting). A single sagittal slice of the resulting Z-statistic image of the first patient is shown in the right panel of Figure 2.3.

In the following analysis, we fit our proposed CWAS model to these Z-statistic images. Bayesian decision theory, using the loss function defined in Section 2.2.3,

is subsequently used to identify activated and deactivated voxels. For comparison, results from the RH and BN models are shown as well (cf. Figures 2.4 and 2.5). We ran our model for 50,000 iterations after a burn-in of 100,000 (total iterations 150,000). The MCMC algorithm takes approximately 1 hour CPU time on an iMAC with a 3.2 GHz Intel Core i5 processor with 16 GB of memory.

Based on our simulation studies, we use a Beta(2,2) prior for the parameters $p_i, i = 1, ..., N$, since this prior showed consistently good performance in identifying activated regions over a large range of thresholds for SNR = 2, 3. Welvaert et al. (2011) demonstrated that the estimated SNR from real fMRI data is around 3.87. Next, we want to determine a threshold to identify activated and deactivated voxels. Since, in pre-surgical fMRI, misclassification of activated or deactivated voxels is more important than false positives, we set $k_1 = 11,\ k_2 = 1,\ t = 1$ in the posterior expected loss function. This results in a loss that penalizes false negatives 11 times more heavily than false positives and was determined by expert opinion (A.J. Bartsch) and is concordant with the results obtained by ESM (see below). We classify any voxel whose posterior expected mean intensity is positive, i.e. $\mathbb{E}(\mu_i \mid \boldsymbol{Y}) > 0$, and for which $f(m_i) = m_i/q_0(\mathbf{m}) \geq 0.2$ as activated. Conversely, those voxels for which the posterior expected mean intensity is negative, i.e. $\mathbb{E}(\mu_i \mid \boldsymbol{Y}) < 0$, and $f(m_i) \geq 0.2$ are considered deactivated.

Figures 2.4 and 2.5 display the activated and deactivated regions, for the two patients, projected onto the cortical surface identified by the CWAS, RH and BN models. Reconstruction of the cortical surface, surface editing to visualize the tumor, transformation/projection of the statistical maps onto the pial surface and the underlying volume space was performed using FreeSurfer v.5.3.0 (http://surfer.nmr.mgh.harvard.edu/; Fischl (2012)) and its recon-all, reg-feat2anat and feat2surf scripts. Left hemispheres are visualized using Freeview (part of FreeSurfer). A minimal amount of smoothing was applied to the pial surface and surface projection of the data, respec-

tively, and is the result of a spline interpolation of the data into the high-resolution ($1.00 \times 1.00 \times 1.00\text{mm}^3$) anatomical volume space. The tumors are made visible through an edited "hole", i.e. a topological defect, of the cortical surface.

Although surgeons are more concerned with activated regions, in Figures 2.4 and 2.5 we also present deactivation regions because in brain tumor patients even decreases in the BOLD signal may be related to brain activation (Fujiwara et al., 2004; Hsu et al., 2004; Ulmer et al., 2004; Bartsch et al., 2006). As a result, deactivated voxels should be considered in the loss function (although we acknowledge that "paradoxical" BOLD signal decreases may often represent a rather slightly, i.e. by the order of a few seconds, increased delay of the peri- or intra-tumoral HRF and are then less likely to reflect true brain activation in fMRI designs of longer blocks such as 30 seconds modeled with temporal derivatives).

Activated and deactivated areas detected by the BN model are clearly less extensive than those obtained by either the CWAS or RH model. While the BN model still performed reasonably in the case of the first patient (Figure 2.4), it clearly did not in the case of the second patient (Figure 2.5). On the other hand, activated and deactivated regions detected by the RH model are obviously larger and "smoother", i.e. more continuously distributed over surface vertices/volume voxels, than those obtained under the CWAS model. Under the RH model, the spatial extent exceeded beyond what is expected according to expert opinion (A.J. Bartsch) in both clinical cases.

Furthermore, these impressions and superiority of the CWAS model were confirmed by intra-operative ESM in both patients—in the first patient, there was fading evidence for activation according to the RH model at the tip of the pars trigangularis of the inferior frontal gyrus (IFG or F3) towards the Sylvian fissure which was eloquent based on electrical stimulation (i.e., ESM evoked speech arrest at this location) and active according to the CWAS results. Compared to the BN model, CWAS re-

sults seemed superior at the peri-Sylvian supramarginal gyrus (SMG) of the inferior parietal lobule (P2) which was not stimulated by ESM in this case but is a known speech center of the dorsal stream, i.e. the Sylvian fissure at the parietotemporal boundary (or Spt) area (cf. Hickok and Poeppel (2007)).

In the second patient, the BN model essentially failed to reveal peri- or intra-tumoral fMRI activations. Here, the RH model suggested high probabilities of activation in the middle temporal gyrus (MTG or T2) which were only confirmed for the area around the superior temporal sulcus (STS or t1) by ESM at approximately the line indicated by the CWAS statistical image thresholded at ≥ 0.20 (cf. lower left panel in Figure 2.5).

Evidence for intra-tumoral fMRI activations was low according to all three models in the first patient and essentially absent on the thresholded CWAS, RN and BH maps (cf. Figure 2.4, lower panel). This was confirmed by ESM which did not lead to speech arrest or other speech and language disturbances during awake craniotomy upon stimulation of the tumor itself. In this case, the resection was limited by proximity to and infiltration of the arcuate fasciculus by the tumor (not shown; cf. Bartsch et al. (2014), with Fig. 23.14 there for illustration and further discussion of eloquent fiber pathways that can be tracked by diffusion-weighted imaging and are similarly important for pre-surgical planning). On the contrary, there was evidence for intra-tumoral activations in the second patient according to the CWAS and RH model which survived the thresholding only in CWAS (cf. Figure 2.5). ESM of the tumor in this case did in fact worsen naming, comprehension and reading performance in the patient and eventually lead to speech arrest which limited the surgical resection. Function in tumor is a known phenomenon to occur primarily in highly malignant tumors (like GBM in the second patient; Ojemann et al. (1996)). Note that the correspondence between ESM and the CWAS, RH and BN maps was assessed after the surgery and was not used for the actual medical decision making.

We conclude this section with a discussion on the quality of the model parameter estimates. Given the large number of parameters in the CWAS model one may wonder whether there is enough information in the data to inform on all of the parameters. The answer is yes. Patient 1's dataset consists of approximately 47000 voxels and patient 2's dataset consists of approximately 54000 voxels. For each voxel, there are three parameters to estimate, μ_i, p_i, and σ_i^2. Although there are many more parameters to estimate than voxels, there is information in the data to estimate all of the parameters. First, the Gelman-Rubin diagnostic statistic shows that the MCMC chains for the μ_is, σ_i^2s, and p_is all have converged. For each parameter, the Gelman-Rubin R statistic (Gelman and Rubin, 1992; Brooks and Gelman, 1998) is calculated from five MCMC simulations with different starting values. The ranges of the Gelman-Rubin diagnostic statistic for the μ_is, p_is, and σ_i^2s are shown in Table 2.2. All ranges are below the suggested threshold of 1.2. Note that we used the univariate Gelman-Rubin diagnostic statistic. Although there is a multivariate version, it is not possible to use it in our application as we have far too many parameters. Figure 2.6 (a)-(c) and Figure 2.7 (a)-(c) display scatterplots of these parameters from two independent MCMC runs from patient 1 and patient 2 respectively. These figures further corroborate the conclusions from the Gelman-Rubin statistic. In Figure 2.6 (d) and Figure 2.7 (d), we also show a histogram of all of the posterior means of the p_is for each patient. The red curves, overlaying the histograms, are the theoretical distribution of the posterior expectation of the p_is under the assumption that the data are completely uninformative and the p_is are independent. It is evident from these two figures that the data are informing the p_is, even though we assume that they are independent a priori. Figure 2.8 shows a histogram of the MCMC standard errors of the μ_is and a scatterplot of MCMC standard errors versus the marginal posterior mean estimates of the μ_is for each patient. MCMC standard errors were calculated using the "mcmcse" package in R (Flegal and Hughes, 2012) with the default batch

means method (Flegal et al., 2008).

2.6 Discussion

In this paper we propose an alternative model to the NP-Potts model developed by Johnson et al. (2013) for analyzing pre-surgical fMRI data. We have developed a novel spatially adaptive smoothing model (CWAS) that allows the data to determine where, and how much smoothing should occur in the Z-statistic image. We have incorporated a Bayesian decision theoretical approach to classify the voxels into activation, deactivation and null states. Simulating from the posterior of the CWAS model is much more computationally efficient than simulating from the non-parametric Potts model proposed by Johnson et al. (2013). The MCMC algorithm for the CWAS model takes one hour to run on an iMac (3.2 GHZ Intel Core i5, 16 GB memory). The computational cost of the NP-Potts model is on the same order as that of the CWAS model only after the ratios of the normalizing constants have been computed. However, it takes more than one week to obtain good estimates of these ratios, rendering the algorithm impractical. Simulation results show that our model outperforms the two existing spatially adaptive smoothing models proposed by Brewer and Nolan (2007) and Reich and Hodges (2008). Finally, our model is easy to implement and converges faster than the other two.

Due to the computational intractability of the true likelihood, we resort to the pseudo-likelihood approximation proposed by Besag (1975) for the joint distribution of $\boldsymbol{\mu}$. Although the pseudo-likelihood approximation solves the problem and Besag (1977) showed that parameter estimation using this approach is efficient for simple Gaussian fields, using the pseudolikelhood approximation brings up some issues. The MCMC samples do not come from the true model and thus the inference is actually based on the likelihood from a misspecified model. It has been shown that the credible intervals based on MCMC samples from pseudo-likelihood have smaller empirical

coverage than the nominal and less than the ones from true likelihood. Shaby (2014) has proposed an adjustment for the MCMC samples, called the *open faced sandwich* (OFS) adjustment, to obtain the correct interval widths and coverage. We looked into the pseudo-likelihood adjustment method described by Shaby (2014) and found that it is impossible to apply it to our results. First, the OFS adjustment doesn't adjust the Bayesian estimators when the loss function used for Bayesian estimation is squared error loss. In this case, the Bayesian estimators are the posterior means and the OFS adjustment cancels out. Thus, the posterior means of the adjusted samples are exactly equal to that of the non-adjusted samples. Hence, under squared error loss, the OFS adjustment only adjusts the width of the credible intervals. In our case, that means that the OFS adjustment only affects the estimates of the m_is since the m_is depend on the variance of the MCMC samples. However, to do the adjustment, we would need to estimate $\Omega = Q^{-1}P^{\frac{1}{2}}Q^{\frac{1}{2}}$, where Ω, Q and P are $n \times n$ matrices, with n number of parameters. Although it is possible to estimate Q and P post hoc, our model has tens of thousands of parameters, thus to derive Ω we need to determine the inverse of a huge matrix, which is computationally intractable. We note that the examples considered by Shaby (2014) all have less than 5 parameters.

In fMRI studies, controlling the family wise error rate (FWER) by restricting to anatomical region-of-interests (ROIs) can increase the statistical sensitivity since the number of multiple comparisons is drastically reduced. However, there is no loss (other than computational) for performing a whole brain analysis. In fact, a whole brain analysis has several advantages. In principle, brain activations are task- or condition-specific. This becomes obvious when you compare the activations in Figures 2.4 and 2.5. The tongue-twister task is phonetically and motorically challenging and therefore more centered to the anterior speech areas involving the precentral gyrus (Figure 2.4), while reading nonfinal embedded clause sentences involves more posterior speech and language areas including the angular gyrus and the so-called

ventral stream of speech and language in general (Figure 2.5). However, contrary to the primary motor or visual system, for example, speech and language have no 'absolute' cortical representations (Bartsch et al., 2014). That is, you can not use anatomical landmarks alone (like the central sulcus to delineate precentral gyrus of the primary motor system) to predict the areas where activations can be expected. In fact, this is reason why we often, and primarily, map speech and language. Furthermore, in pre-surgical fMRI of patients with brain lesions, compensatory mechanisms and 'neuroplasticity' may lead to atypical and unexpected activations (at the border or even outside anatomical ROIs defined by task-specific functional activations evoked in a group of healthy comparison subjects, for example) which we don't want to miss. Last but not least, performing a whole-brain analysis instead of an analysis restricted to predefined anatomical ROIs (as derived from prior functional studies, for example) or even just the intratumoral and peritumoral regions and the potential routes of neurosurgical access allows us to assess the general pattern of activation and deactivation, with both the expected as well as unexpected components (Figures 2.4 and 2.5), which proves very helpful to assure that the experiment worked (even in brain areas remote from the tumor). This is also why almost all pre-surgical fMRI studies record the data from the entire brain. Taken together, the whole-brain approach does not involve a statistical penalty for our analysis but provides us with a much richer set of information.

We applied the CWAS model to Z-statistics images from a mass univariate GLM approach. However, our model can be applied to the statistical output of independent component analyses (ICA) such as FSL's MELODIC (Beckmann and Smith, 2004). For speech and language mapping by fMRI, ICA may be more appropriate because it does not assume that brain activation is maintained at a similar level during one block (which is quite unlikely given the rather lengthy blocks of about 30 seconds in clinical fMRI).

A further extension to the CWAS model would be to modify the priors on the p_is so that they are explicitly spatially dependent a priori. In this paper, the prior distribution of the c_is is induced by the prior distribution on the p_is and the p_is/c_is are independent since $p_i \overset{iid}{\sim} \text{Beta}(\alpha, \beta), i = 1, \ldots, N$. If one wished to model the p_is as spatially dependent, while retaining control on the degree of smoothing, a prior which combines the feature of a CAR prior and a Beta distribution could be employed:

$$\pi(p_i) \propto \exp\left\{-0.5 \sum\nolimits_{i \sim j} (p_i - p_j)^2/\varphi^2 + (\alpha - 1) \log p_i + (\beta - 1) \log(1 - p_i)\right\}.$$

Another interesting extension of the CWAS model would be to modify it to accommodate multiple fMRI studies from the same patient, or multiple runs using the same paradigm. Consider, for example, a sentence completion experiment being performed in addition to a tongue-twister experiment. If there is any correlation between the two Z-statistic images, a bivariate CWAS model may be able to leverage it and produce more accurate results. Furthermore, instead of using it across different runs with different paradigms such a model may especially be helpful in analyzing multiple runs using the same paradigm where, normally, a second level fixed-effects analysis is conducted. We are currently investigating several multivariate spatially adaptive CAR models for such analyses.

Center Voxel	[20,40,10]	[36,50,18]	[31,35,20]	[53,29,25]	[40,40,30]	[46,25,33]
Effect Size	10	8	6	4	2	1
Radius	2	4	1	2	4	1

Table 2.1: Locations, effect sizes and radii of the 6 activated regions in the simulation studies.

	μ_is	p_is	σ_i^2s
Patient 1	(0.999952, 1.000414)	(0.999954, 1.027688)	(0.999953, 1.000770)
Patient 2	(0.999953, 1.000486)	(0.999950, 1.000052)	(0.999950, 1.000466)

Table 2.2: The ranges of the Gelman-Rubin diagnostic statistic for the μ_is, p_is, and σ_i^2s.

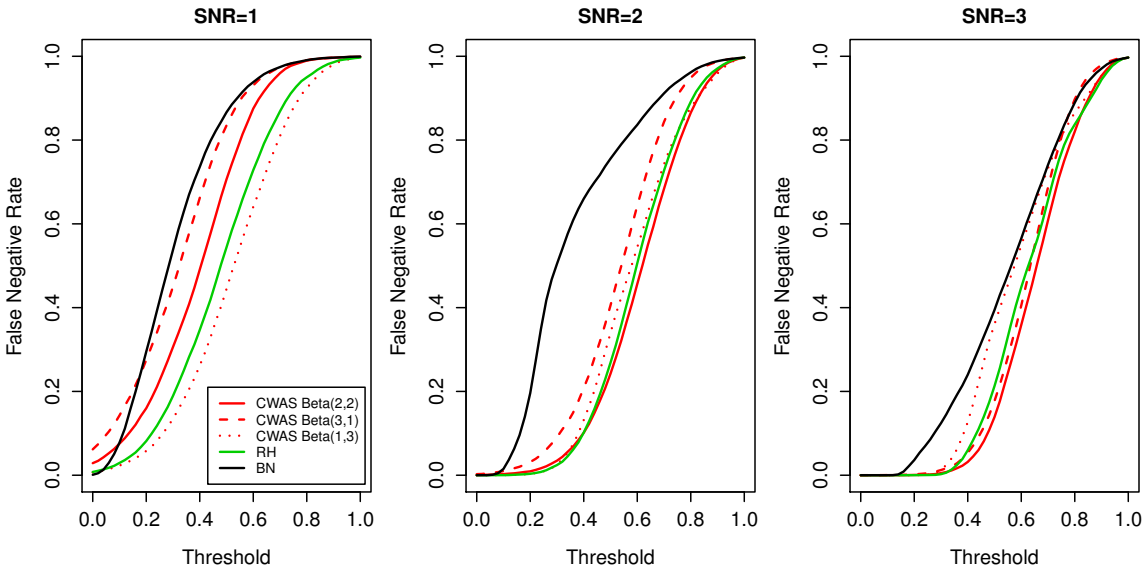

Figure 2.1: False negative rate vs threshold in the simulation studies. For each signal to noise ratio, we take the average of the false negative rate over 50 simulations using different Beta prior distributions for the p_is in our CWAS model. For comparison we also include results using the BN and RH models. Threshold on the loss function scale $f(m_i)$ varies from 0 to 1.

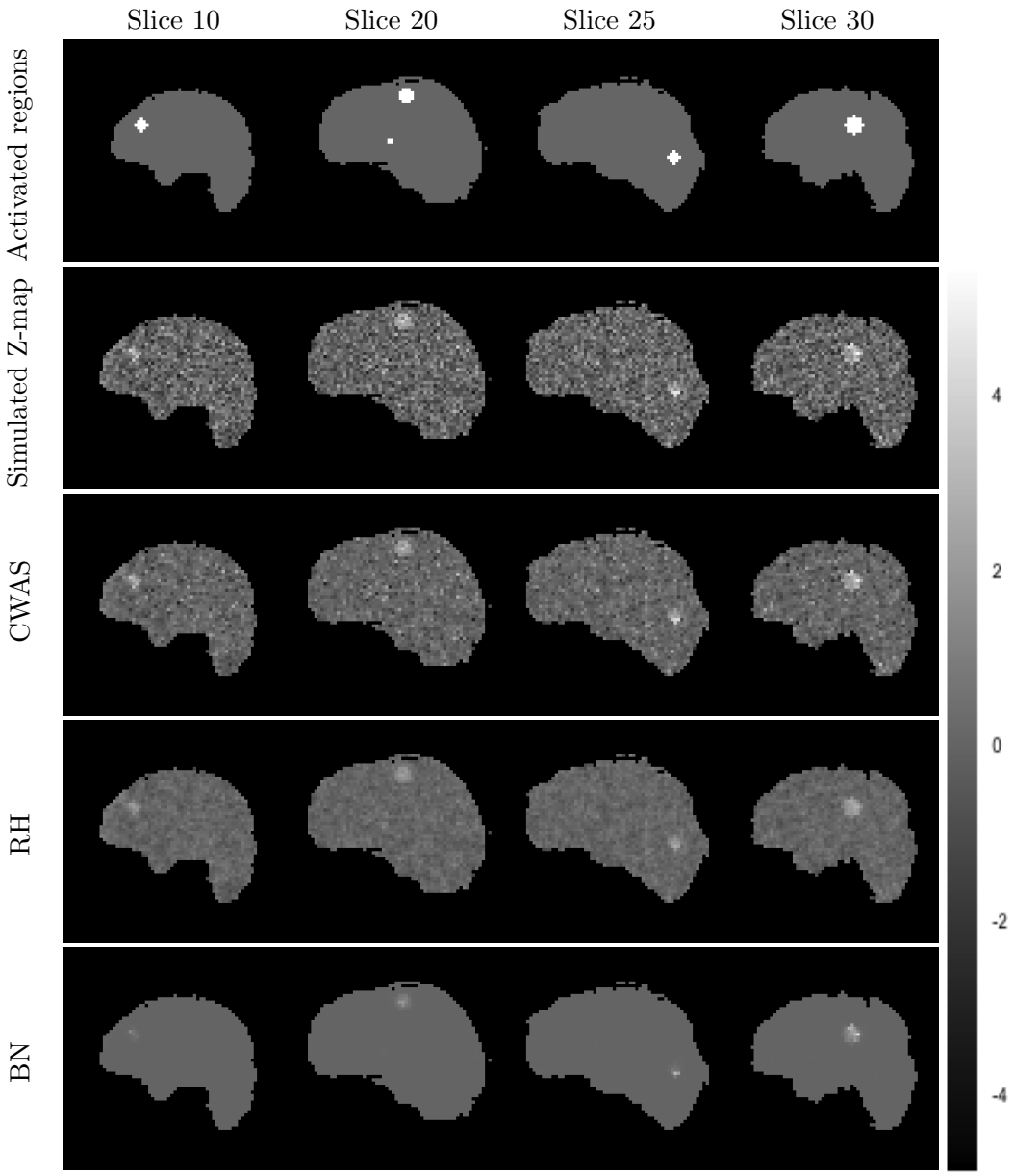

Figure 2.2: Comparison of the posterior estimates of the mean intensities from the three different spatially adaptive smoothing models when SNR = 2. The top row shows the truly activated regions in four sagittal slices of the brain (one of the six simulated regions does not appear in these four slices). The second row shows the simulated Z-statistic images (with no smoothing). The third through fifth rows show the marginal posterior means of $\boldsymbol{\mu}$ from the three spatially adaptive smoothing models.

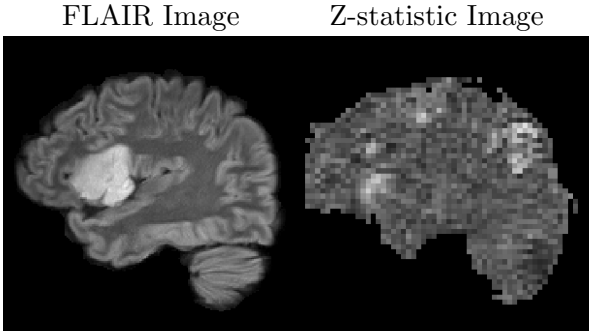

Figure 2.3: Pre-surgical fMRI imaging. The left panel shows a FLAIR (fluid-attenuated inversion recovery) image where the tumor of the first patient is evident. The right panel shows the unsmoothed Z-statistic image. The Z-statistic image generated by mass-univariate general linear modeling is related to the probability of brain activation.

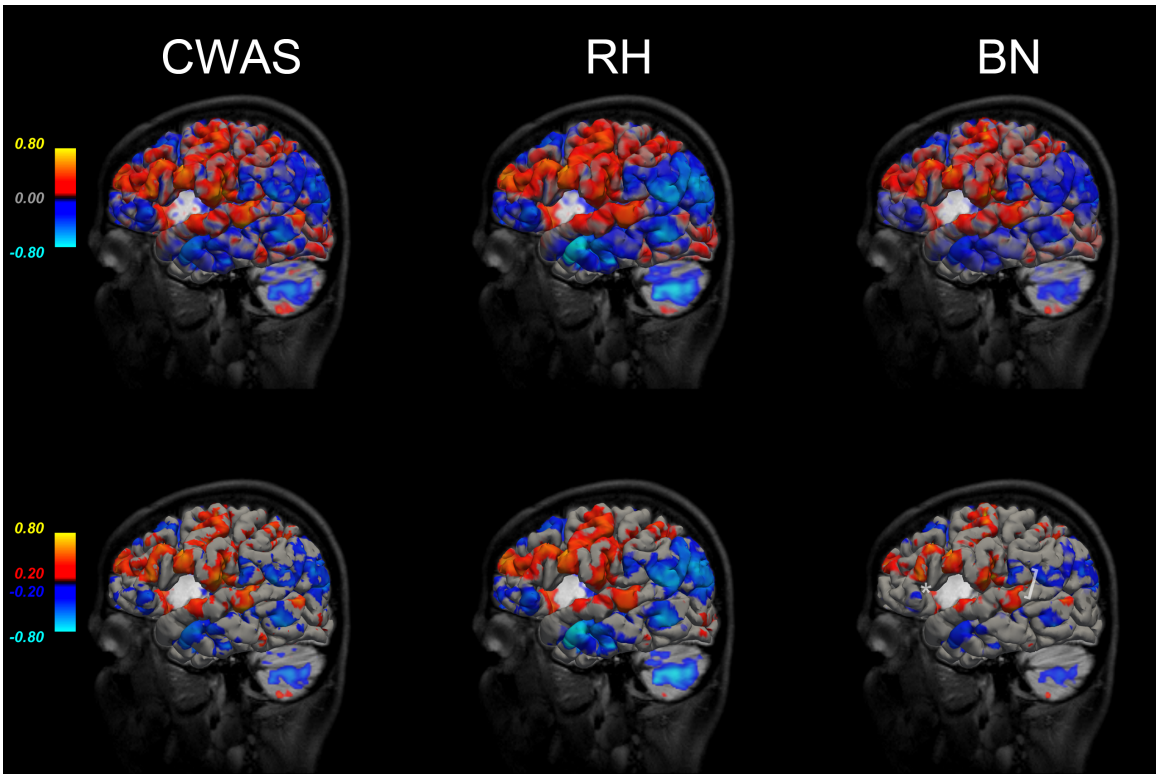

Figure 2.4: Results for patient 1: CWAS, RH and BN models. The top row shows results on the loss function scale projected onto the cortical surface. In the bottom row, the loss function has been truncated below by 0.2 and above by -0.2—indicating those regions that survived thresholding. Values above 0.8 are mapped to 0.8 and values below -0.8 are mapped to -0.8 to give a better dynamic range of colors. Red to yellow denote increasingly strong activation. Dark blue to bright blue denote increasingly strong deactivation. In the lower right panel, the tip of the pars triangularis of the inferior frontal gyrus is marked with a white asterisk * and the peri-Sylvian supramarginal gyrus (SMG or area Spt) is marked with a slightly rotated white bracket].

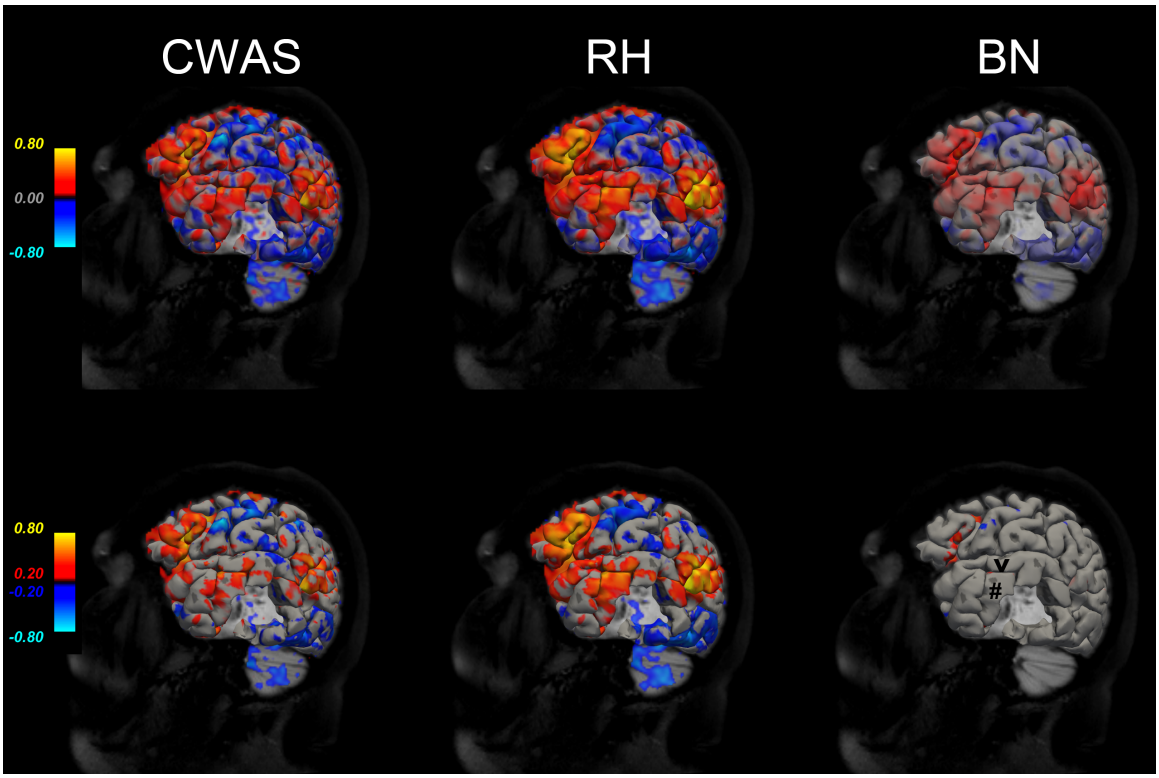

Figure 2.5: Results of the CWAS, RH and BN models for patient 2 (cf. Figure 2.4 for further explanations). In the lower right panel, the superior temporal sulcus (STS or t1) is marked with a black arrowhead and the middle temporal gyrus (MTG or T2) is marked with a black #.

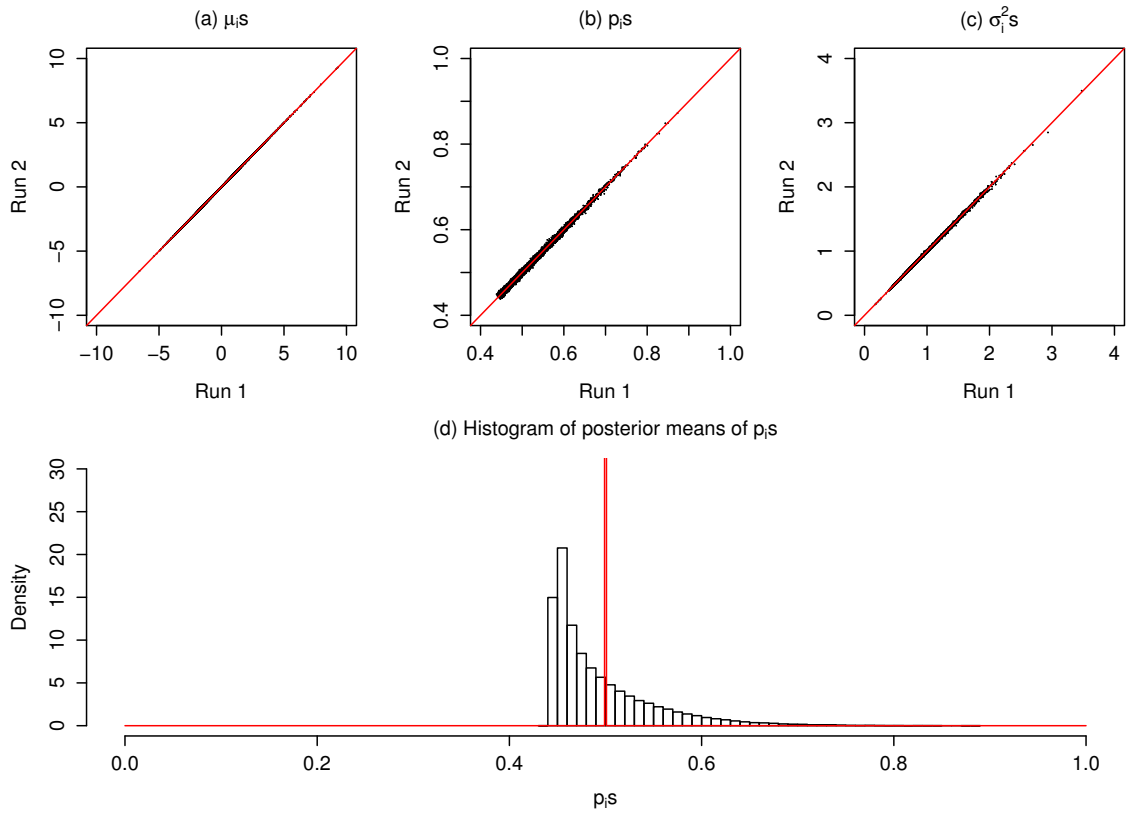

Figure 2.6: (a)-(c): Scatterplots of posterior mean estimates from two independent MCMC runs from patient 1. (d): Histogram of the posterior estimates of p_is with red line indicating the sampling distribution of the posterior means of p_is from patient 1 run 1.

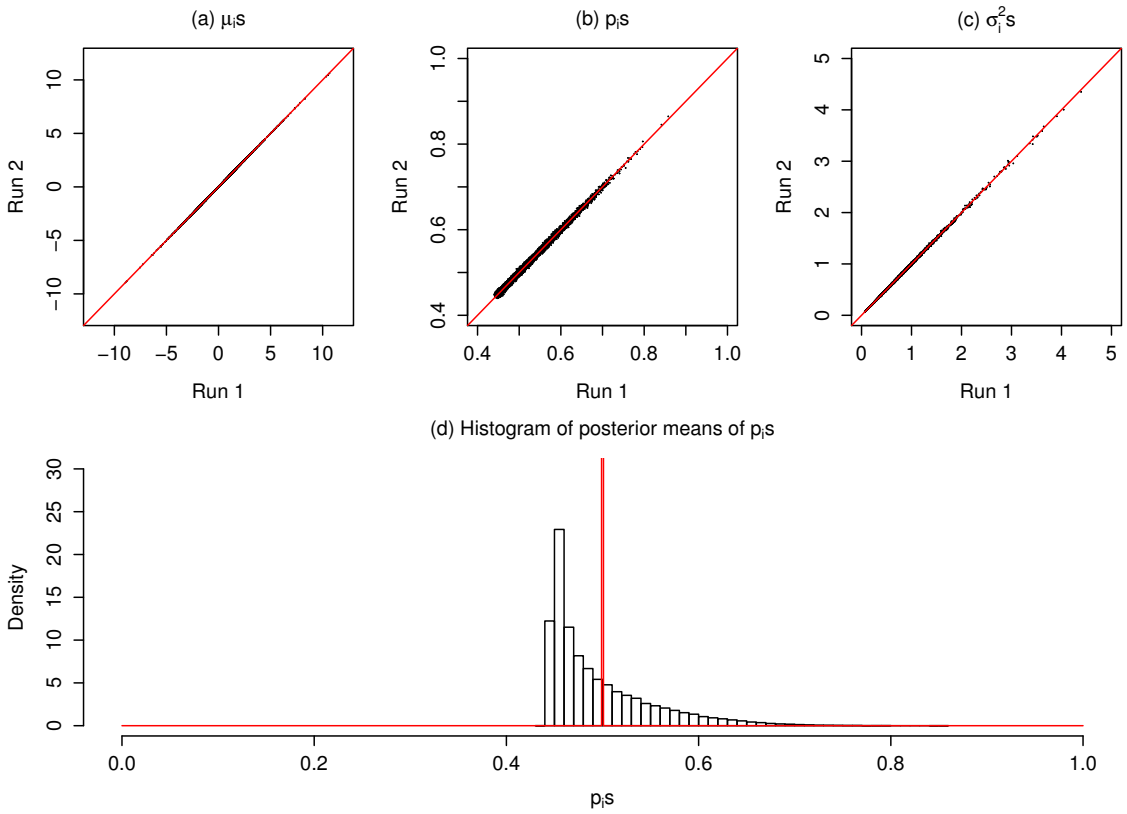

Figure 2.7: (a)-(c): Scatterplots of posterior mean estimates from two independent MCMC runs from patient 2. (d): Histogram of the posterior estimates of p_is with red line indicating the sampling distribution of the posterior means of p_is from patient 2 run 1.

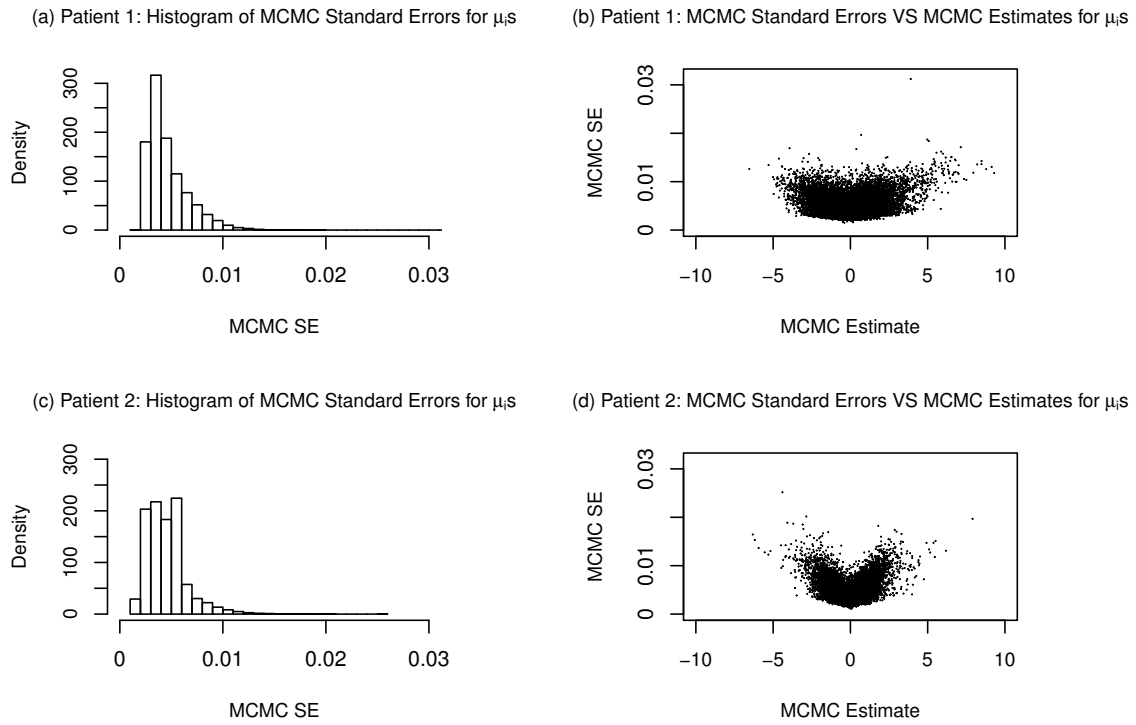

Figure 2.8: MCMC standard errors (SE) for the μ_is. Left column: histograms of the MCMC standard errors for the μ_is for patient 1 (a) and patient 2 (c). Right column: scatterplots of MCMC standard errors versus marginal posterior means of the μ_is for patient 1 (b) and patient 2 (d).

CHAPTER III

Multivariate Spatially Adaptive Conditionally Autoregressive Models with Application to Multivariate FMRI Data

3.1 Introduction

As noted in Chapter I, spatial smoothing is a necessary preprocessing step for fMRI data analysis: it increases the signal to noise ratio and it ensures that the data follows the assumptions of Gaussian random field theory, which is required in the subsequent multiple comparisons procedure (Worsley et al., 1996). Typically, in fMRI data analysis spatial smoothing is achieved by convolving the data with a Gaussian kernel of fixed width, with the latter being identified by the full width of the kernel at half of its maximum height (FWHM). This spatial smoothing is commonly implemented in software such as SPM (Statistical Parametric Mapping, Friston et al. (1994a)) and FSL (fMRIB Software Library, Smith et al. (2004); Woolrich et al. (2009)). Although widely adopted, smoothing fMRI data using a fixed FWHM across the entire image has several drawbacks, mostly related to potential under- and over-smoothing of brain images. This in turn can lead to incorrect identification of activated/deactivated regions in the brain. The consequences of such misclassification might be severe, especially in a pre-surgical fMRI analysis context. To alleviate this

problem, in Chapter II, we presented a conditionally weighted adaptive smoothing model (CWAS) that allowed the amount of smoothing to vary across the image. In its current formulation, the CWAS model can only be applied to one fMRI image at a time, while it is often the case that multiple fMRI images are available from the same patient, either because the patient has been exposed to a variety of experimental conditions or because the patient has been subject to the same experiment several times. To leverage the information contained in the multiple brain images pertaining to the same subject, in this chapter, we present a series of models to extend the idea of spatially adaptive smoothing to multiple pre-surgical fMRI images. These models, which will be specified locally, can be thought as multivariate spatially adaptive CAR models which exploit local correlations within and across images to improve statistical inference and identification of activated and deactivated regions.

Building upon our previous work, the simplest strategy to adaptively smooth multiple fMRI datasets would be to use the CWAS model on each image separately. However, this approach would not allow us to leverage the correlation between images. Therefore, our intent here is to propose a multivariate extension of the CWAS model that explicitly accounts for the correlation across images. Recall that the univariate CWAS model, which we use as a prior for the smoothed voxel intensities (see Chapter II), is a modification of the univariate CAR model where the variances of the full conditionals are expressed as a proportion of the unsmoothed voxel intensity variances, and not as a function of a global variance parameter, as well as the degree of connectivity of a voxel. Thus, in extending the CWAS model to a multivariate setting, one approach is to slightly modify the multivariate CAR model.

Different approaches have been proposed in the literature to extend the univariate CAR model to a multivariate setting. In an analysis of ecological epidemiological data, Kim et al. (2001) proposed a bivariate CAR model ($p = 2$) to model the mortality rates of a disease while incorporating the information from a second disease.

This model that the authors called a "twofold CAR model" is hard to generalize to $p \geq 2$. Building upon the family of models introduced by Mardia (1988), Carlin and Banerjee (2003) and Gelfand and Vounatsou (2003) developed multivariate CAR (MCAR) models that include spatial autoregressive parameters to ensure that the multivariate distribution is proper. Finally, Jin et al. (2005) proposed the generalized MCAR model that directly specifies the joint multivariate distribution through the simpler conditional and marginal distributions. Despite the different nuances of each of the above mentioned MCAR models, they all assume that the spatial correlation within images depends on a global covariance parameter. In this study, we relax the assumption of a global covariance/correlation parameter, and analogously to the univariate CWAS model, we allow the covariance matrices accounting for the correlation between images to vary spatially.

Also in the context of fMRI data analysis, there have been efforts to integrate multiple fMRI datasets. For instance, fixed-effects and mixed-effects analyses are two types of approaches implemented when dealing with multiple fMRI datasets (Penny et al., 2003). Single-subject analyses are generally carried out with the fixed-effects approach since this approach only takes into account within-subject (also known as between-scan) variability. Mixed-effects analysis, by contrast, takes into account both within-subject and between-subject variabilities, and thus allows inferences to be made about the population from which the subjects are drawn. Mixed-effects analyses are typically performed with a two-level summary statistic procedure proposed by Holmes and Friston (1998) which has the advantage of simplicity and computational feasibility. In a nutshell, the approach of Holmes and Friston (1998) consists of a two-step analysis. First, different general linear models (GLM) are independently fitted for each subject, which yields first level inference statistics. In the second level, the summary statistics from the first-level analyses are inputs for the second-level GLM to account for between-subject variability. Still using a two-level summary

statistics framework, Friston et al. (2002) and Beckmann et al. (2003) formulate their approach in a hierarchical fashion. Thus, they proposed a hierarchical model where the different levels account for session, subject and/or group, respectively. Using the same hierarchical formulation, Woolrich et al. (2004) proposed inference techniques on the top level for hierarchical models with multiple levels in a Bayesian framework. One key problem associated with the hierarchical and two-stage level summary statistics approaches described above, is the lack of clarity on which summary statistics should be used to ensure the validity of inference at the top level, as pointed out by Woolrich et al. (2009). Furthermore, statistical inference for all these approaches are based on mass univariate GLMs, which rely on the assumption that intensities at voxels are spatially independent.

In this chapter, building upon the spatial statistical literature on MCAR models, we propose three spatially adaptive multivariate conditional autoregressive models, which we call SAMCAR. The first model extends the MCAR model of Carlin and Banerjee (2003) and Gelfand and Vounatsou (2003) and allows the covariance matrices in the full conditionals to spatially vary. The second model is a direct extension of the univariate CWAS model to accommodate multiple fMRI datasets, while the third model is a multivariate generalization of the spatially adaptive CAR model proposed by Reich and Hodges (2008), described in Section 2.2.2.2 and named the RH model. We compare the performance of these models to that of a mixed-effects model with spatially varying coefficients to account for the spatial correlations between and within fMRI datasets. All the models considered, the three SAMCAR models and the mixed-effects model, not only allow us to smooth multiple fMRI images in a spatially adaptive way, but also allow us to better quantify the spatial correlations among multiple fMRI images.

The remainder of this chapter is organized as follows. In Section 3.2, we describe the original MCAR model along with the three spatially adaptive multivariate con-

ditional autoregressive models that we propose as well as the mixed-effects model. In Section 3.3, we conduct two simulation studies to compare the performance of the MCAR model, with that of the three proposed SAMCAR models, the mixed-effects model and the CWAS model applied to each dataset separately. In Section 3.4, we consider real data applications and using two different sets of two fMRI Z-statistic images. The fMRI datasets are acquired either from the same patient using the same paradigm or from the same patient but with different paradigms. In each analysis, our goal is to identify task-related brain activated and deactivated regions. We conclude the chapter with a brief discussion in Section 3.5.

3.2 Models

In this section, we first review the formulation of the MCAR model, then introduce the three proposed SAMCAR models (spatially adaptive multivariate CAR models), that we developed building upon the MCAR model, the CWAS model and the RH model, respectively. Lastly, we describe the proposed mixed-effects model that we use to integrate multiple fMRI datasets.

3.2.1 The MCAR Model

Consider K unsmoothed 3-dimensional Z-statistic images obtained from fMRI experiments conducted either on the same subject or on multiple patients. To combine the data across different sessions or subjects, we first register all the brain images into a common anatomical space. After registering every dataset, the resulting data consist of K Z-statistic images that all have exactly the same shape and the same number N of voxels. As voxels are 3-dimensional volumes in the brain, within the brain, we can define a first order neighborhood system where two voxels, i and j, are called neighbors if they share a common face. Mathematically, we denote the adjacency relationship between voxel i and j as $i \sim j$. Let Y_{ik} be the observed Z-statistic

intensity relative to the ith voxel in the kth image, and $\boldsymbol{Y}_i = (Y_{i1}, Y_{i2}, ..., Y_{iK})^T, i = 1, 2, ..., N$ be the vector of intensities for voxel i across the K images. Finally, let $\boldsymbol{Y} = (\boldsymbol{Y}_1^T, \boldsymbol{Y}_2^T, ..., \boldsymbol{Y}_N^T)^T$ be a vector of dimension NK that include all the voxel intensities obtained in K images. We postulate that at any given voxel in any given image the observed intensity can be decomposed into the sum of the true mean intensity at the voxel plus some measurement error:

$$y_{ik} = \mu_{ik} + \epsilon_{ik},\ \epsilon_{ik} \overset{iid}{\sim} \mathrm{N}(0, \tau^2),\ i = 1, 2, ..., N,\ k = 1, ..., K. \tag{3.1}$$

In other words, we assume that the variance of the measurement error is constant across images and across voxels.

The MCAR model was formally developed by Carlin and Banerjee (2003) and Gelfand and Vounatsou (2003) building upon the work of Mardia (1988) who presented a full theoretical development of multivariate Gaussian Markov random fields. Analogous to the univariate specification of the CAR model offered in Equation 2.1, the MCAR model is defined through a set of full conditionals; specifically:

$$\left[\boldsymbol{\mu}_i \mid \boldsymbol{\mu}_{(-i)}, \boldsymbol{\Sigma}_i\right] \sim \mathrm{MVN}\left(\boldsymbol{A}_i \sum\nolimits_j \boldsymbol{B}_{ij}\boldsymbol{\mu}_j, \boldsymbol{\Sigma}_i\right), i = 1, ...N, \tag{3.2}$$

where $\boldsymbol{A}_i, \boldsymbol{B}_{ij}, \boldsymbol{\Sigma}_i$ are $K \times K$ matrices.

By Brook's lemma (Brook, 1964) and under regularity conditions, Mardia (1988) shows that the full conditionals in Equation 3.2 uniquely determine the following joint distribution:

$$\boldsymbol{\mu} \sim \mathrm{MVN}\left(0, [\boldsymbol{\Gamma}^{-1}(\boldsymbol{I} - \boldsymbol{B}_{\boldsymbol{A}})]^{-1}\right), \tag{3.3}$$

where $\boldsymbol{\Gamma}$ is a $NK \times NK$ block diagonal matrix with $K \times K$ diagonal entries $\boldsymbol{\Sigma}_i$, $\boldsymbol{I}$ is a NK-dimensional identity matrix, and $\boldsymbol{B}_{\boldsymbol{A}}$ is a $NK \times NK$ matrix with (i, j)th block $\boldsymbol{A}_i\boldsymbol{B}_{ij}$ and (i, i)th block zero.

The joint distribution in Equation 3.3 is proper if and only if $\boldsymbol{\Gamma}^{-1}(\boldsymbol{I} - \boldsymbol{B_A})$ is a symmetric and positive definite (SPD) matrix. In order to develop a model for which assessing the SPD condition is a simple task, Carlin and Banerjee (2003) and Gelfand and Vounatsou (2003) made several simplifying assumptions on the formulation of $\boldsymbol{\Gamma}^{-1}(\boldsymbol{I} - \boldsymbol{B_A})$, all derived starting from the adjacency matrix $\mathbf{W}$, that is, the matrix that reflects the neighborhood system in the spatial domain of interest. More clearly, $\boldsymbol{W}$ is the adjacency matrix with entries w_{ij}, the binary adjacency indicators, equal to 1 if $i \sim j$, and 0 otherwise; $w_{i+} = \sum_{j=1}^{N} w_{ij}$ and $\boldsymbol{D} = \text{Diag}(w_{i+})$ is an N by N diagonal matrix. Carlin and Banerjee (2003) and Gelfand and Vounatsou (2003) introduce the scaled adjacency matrix $\boldsymbol{B}$ defined by $\boldsymbol{B} = \boldsymbol{D}^{-1}\boldsymbol{W}$. Assuming, in Equation 3.3, that $\boldsymbol{\Sigma}_i = \boldsymbol{\Lambda}/w_{i+}$ where $\boldsymbol{\Lambda}$ is a $K \times K$ positive definite and symmetric matrix, $\boldsymbol{A}_i = \alpha \boldsymbol{I}_{K\times K}$, with $\boldsymbol{I}_{K\times K}$ K-dimensional identity matrix, and $\boldsymbol{B}_{ij} = b_{ij}\boldsymbol{I}_{K\times K}$ where b_{ij} is the (i, j) entry of matrix $\boldsymbol{B}$, Equation 3.3 is simplified to:

$$\boldsymbol{\mu} \sim \text{MVN}\left(0, [(\boldsymbol{D} - \alpha \boldsymbol{W}) \otimes \boldsymbol{\Lambda}^{-1}]^{-1}\right), \tag{3.4}$$

with $\otimes$ denoting the Kronecker product.

Under these assumptions, the symmetry condition of $\boldsymbol{\Gamma}^{-1}(\boldsymbol{I} - \boldsymbol{B_A})$ is satisfied. In addition, it can be shown that the positive definiteness condition is satisfied when $\alpha \in (-1, 1)$ (Carlin and Banerjee, 2003; Gelfand and Vounatsou, 2003). For $\alpha = 1$, the distribution is improper since $\boldsymbol{D} - \boldsymbol{W}$ is singular. In our analysis of multiple fMRI images, we use the MCAR formulation given in Equation 3.4 with $\alpha = 1$ to model the true, unobserved smoothed intensities $\boldsymbol{\mu}$. Despite the fact that $\alpha = 1$ leads to an improper joint distribution, we select $\alpha = 1$ because it has been noted that an MCAR model with α not close to 1 does not deliver sufficient spatial smoothing (Carlin and Banerjee, 2003), which is one of our primary goals here. Hence, we use the MCAR model in Equation 3.4 with $\alpha = 1$ as a prior distribution for $\boldsymbol{\mu} = (\boldsymbol{\mu}_1^T, \boldsymbol{\mu}_2^T, ..., \boldsymbol{\mu}_N^T)^T$,

where each $\boldsymbol{\mu}_i$ is an K-dimensional vector, i.e., $\boldsymbol{\mu}_i = (\mu_{i1}, \mu_{i2}..., \mu_{iK})^T$; specifically:

$$\left[\boldsymbol{\mu}_i \mid \boldsymbol{\mu}_{(-i)}, \boldsymbol{\Lambda}\right] \sim \text{MVN}\left(\sum\nolimits_j (w_{ij}/w_{i+})\boldsymbol{\mu}_j, \boldsymbol{\Lambda}/w_{i+}\right), i = 1, ...N. \tag{3.5}$$

We note that the fact that the joint distribution implied by the MCAR model above is improper, this does not constitute a problem since it has been shown in the literature that the posterior distribution is proper (Gelfand and Vounatsou, 2003).

3.2.2 Spatially Adaptive MCAR Models

3.2.2.1 SAMCAR I

As can be noted by the specification above, the MCAR model introduced in Section 3.2.1 assumes that the covariance between the intensity of a voxel in an image and the intensities of the same voxel in the other images can be described by a general $K \times K$ covariance matrix $\boldsymbol{\Lambda}$ appropriately rescaled to account for the connectivity level of the voxel, as the conditional covariance matrix for each voxel in Equation 3.5, $\boldsymbol{\Lambda}/w_{i+}$, shows. For fMRI data, such assumptions lead to extra variability at the edges of the brain and a constant level of variability elsewhere, since voxels on the edge of the brain have less than six neighbors while internal voxels have six neighbors. To relax the assumption of a common covariance matrix $\boldsymbol{\Lambda}$, we propose a spatially adaptive MCAR model with a different conditional covariance matrix $\boldsymbol{\Omega}_i$ for each voxel i, that is:

$$\left[\boldsymbol{\mu}_i \mid \boldsymbol{\mu}_{(-i)}, \boldsymbol{\Omega}_i\right] \sim \text{MVN}\left(\sum\nolimits_j \frac{w_{ij}}{w_{i+}}\boldsymbol{\mu}_j, \boldsymbol{\Omega}_i\right), i = 1, ..., N. \tag{3.6}$$

The diagonal elements of $\boldsymbol{\Omega}_i$ account for the variability in intensity at voxel i within images while the off diagonal elements account for the variability in intensity at voxel i between images. In turn, the matrix $\boldsymbol{\Omega}_i$ is provided with a prior which, without loss

of generality, in what follows we describe the case of $K = 2$. More specifically, we assign an inverse Wishart prior with scale matrix a 2×2 (in general, $K \times K$) identity matrix and degrees of freedom equal to 2, i.e., $\boldsymbol{\Omega}_i \overset{iid}{\sim} \text{IW}(\boldsymbol{I}_{2\times 2}, 2)$. Note that the degrees of freedom can be considered as the a priori sample size, and thus a smaller value of the degrees of freedom implies a rather uninformative prior.

The first level of this model, which we call the spatially adaptive MCAR Model I (SAMCAR I), is specified the same as Equation 3.1: $y_{ik} = \mu_{ik} + \epsilon_{ik}$, $i = 1, 2, ..., N$, $k = 1, 2$, where $\epsilon_{ik} \overset{iid}{\sim} \text{N}(0, \tau^2)$. To complete the prior specifications of this model, we place an inverse Gamma distribution on the measurement error variance τ^2: $\tau^2 \sim \text{IG}(1, 1)$. Again with a large number N of voxels, the priors on τ^2, as the prior on $\boldsymbol{\Omega}_i$, $i = 1, \ldots, N$ is relatively weak.

3.2.2.2 SAMCAR II

The second spatially adaptive MCAR model (SAMCAR II) that we propose is an extension of the conditionally weighted adaptive smoothing (CWAS) model developed in Chapter II. Referring back to Equation 3.2, the SAMCAR II model is defined by assuming a compound symmetric covariance structure for the conditional variance $\boldsymbol{\Sigma}_i$. In other words, we assume the following specification for the full conditionals on the true, smoothed intensity $\boldsymbol{\mu_i}$, $i = 1, \ldots, N$:

$$\left[\boldsymbol{\mu}_i \mid \boldsymbol{\mu}_{(-i)}, \sigma_i^2, \rho_i\right] \sim \text{MVN}\left(\sum\nolimits_j \frac{w_{ij}}{w_{i+}}\boldsymbol{\mu}_j, \begin{pmatrix} \sigma_i^2 & \rho_i\sigma_i^2 \\ \rho_i\sigma_i^2 & \sigma_i^2 \end{pmatrix}\right), i = 1, ..., N. \tag{3.7}$$

In this model, instead of using a global measurement error variance τ^2 for all voxels, we let the measurement error vary across the brain and model τ^2 to be proportional to the variance σ_i^2, that is: $y_{ik} = \mu_{ik} + \epsilon_{ik}$, $\epsilon_{ik} \sim \text{N}\left(0, c_i\sigma_i^2\right), i = 1, ..., N, k = 1, 2$, $c_i > 0$. Analogous to the univariate case, the prior distribution of c_i is induced by the prior distribution on p_i. That is, we define c_i as $c_i = (1 - p_i)/p_i$ and we specify

a Beta prior on p_i, e.g. $p_i \overset{iid}{\sim} \text{Beta}\,(\alpha, \beta)\,, i = 1, ..., N$. Based on the simulation results obtained in Chapter II, we set $\alpha = \beta = 2$. As in the CWAS model, to borrow information from the neighbors, we let the variances σ_i^2 vary spatially and we place a CAR prior on their logarithm; that is:

$$[\ln(\sigma_i^2) \mid \ln(\boldsymbol{\sigma}_{(-i)}^2), \lambda^2] \sim \text{N}\left(\sum\nolimits_j (w_{ij}/w_{i+}) \ln(\sigma_j^2),\ \lambda^2/w_{i+}\right), i = 1, ..., N. \tag{3.8}$$

Similarly, as we imagine that the correlations between voxels might vary smoothly across the brain, to allow for spatial smoothing of the correlations, we place a CAR prior on the correlations ρ_is, restricted to the interval $(-1; 1)$:

$$[\rho_i \mid \boldsymbol{\rho}_{(-i)}, \theta^2] \sim \text{N}\left(\sum\nolimits_j (w_{ij}/w_{i+})\rho_j,\ \theta^2/w_{i+}\right) 1(-1 \leq \rho_i \leq 1), i = 1, ..., N. \tag{3.9}$$

3.2.2.3 SAMCAR III

The third spatially adaptive MCAR model (SAMCAR III) that we propose can be considered as a multivariate generalization of the spatially adaptive CAR model proposed by Reich and Hodges (2008), which is described in Chapter II. In the first level, we assume that $y_{ik} = \mu_{ik} + \epsilon_{ik}$, $i = 1, 2, ..., N$, $k = 1, 2$, where $\epsilon_{ik} \overset{iid}{\sim} \text{N}(0, \tau^2)$. Differently from two spatially adaptive MCAR models (SAMCAR I & II) presented earlier, whose full conditional specification did not lead to a proper joint distribution due to the non-symmetric precision matrix $\boldsymbol{\Gamma}^{-1}(\boldsymbol{I} - \boldsymbol{B_A})$, the SAMCAR III is formulated so as to yield a valid covariance matrix, and thus a proper joint distribution for $\boldsymbol{\mu}$. To allow the correlation between images to vary spatially while maintaining a valid covariance matrix in the joint distribution, to each voxel we associate two

parameters, a variance parameter σ_i^2 and a correlation parameter ρ_i, we define:

$$\boldsymbol{B}_{ij} = w_{ij} \begin{pmatrix} \sigma_i\sigma_j & \rho_i\rho_j\sigma_i\sigma_j \\ \rho_i\rho_j\sigma_i\sigma_j & \sigma_i\sigma_j \end{pmatrix}^{-1}.$$

We then let the matrices $\boldsymbol{\Sigma}_i$ and $\boldsymbol{A}_i$ of Equation 3.2 be equal to $\boldsymbol{\Sigma}_i = \left[\sum_j \boldsymbol{B}_{ij}\right]^{-1}$ and $\boldsymbol{A}_i = \boldsymbol{\Sigma}_i$, respectively. Under these choices, the precision matrix in Equation 3.3 becomes:

$$\boldsymbol{\Gamma}^{-1}(\boldsymbol{I} - \boldsymbol{B}_{\boldsymbol{A}}) = \text{Diag}(\boldsymbol{\Sigma}_i^{-1})(\boldsymbol{I} - \boldsymbol{B}_{\boldsymbol{A}}) = \text{Diag}(\textstyle\sum_j \boldsymbol{B}_{ij}) - \tilde{\boldsymbol{B}},$$

where $\tilde{\boldsymbol{B}}$ is an $NK \times NK$ matrix with (i,j)th entry $\boldsymbol{B}_{ij}$. Combining everything together, our third spatially adaptive MCAR model is defined through the following conditional distributions:

$$\left[\boldsymbol{\mu}_i \mid \boldsymbol{\mu}_{(-i)}, \tilde{\boldsymbol{B}}\right] \sim \text{MVN}\left(\left[\textstyle\sum_j \boldsymbol{B}_{ij}\right]^{-1} \textstyle\sum_j \boldsymbol{B}_{ij}\boldsymbol{\mu}_j, \left[\textstyle\sum_j \boldsymbol{B}_{ij}\right]^{-1}\right), i = 1, ..., N. \tag{3.10}$$

These full conditionals imply the following joint distribution for $\boldsymbol{\mu}$:

$$\pi(\boldsymbol{\mu} \mid \tilde{\boldsymbol{B}}) \propto |\boldsymbol{Q}|^{\frac{1}{2}} \exp\left\{-\frac{1}{2}\textstyle\sum_{i\sim j}(\boldsymbol{\mu}_i - \boldsymbol{\mu}_j)^T \boldsymbol{B}_{ij}(\boldsymbol{\mu}_i - \boldsymbol{\mu}_j)\right\}, \tag{3.11}$$

where $\boldsymbol{Q} = \text{Diag}(\sum_j \boldsymbol{B}_{ij}) - \tilde{\boldsymbol{B}}$ and $|\boldsymbol{Q}|$ is the product of all positive eigenvalues of $\boldsymbol{Q}$.

Finally, we assume $\tau^2 \sim \text{IG}(1,1)$ and we allow σ_i^2 and ρ_i to vary spatially by using the same priors as Equations 3.8 and 3.9.

We note that in the application of SAMCAR III, the computation of $|\boldsymbol{Q}|$ is intractable for the joint distributions of $\boldsymbol{\mu}$ due to the large data size ($N \approx 47000$). For SAMCAR I and II, the joint distributions of $\boldsymbol{\mu}$ are not proper. To overcome these issues, we use the pseudo-likelihood approach (Besag, 1975) to approximate the joint

distributions for $\boldsymbol{\mu}$ for all three models. Specifically, the approximation is formulated as the product of all full conditionals (See Chapter II Section 2.3 for more details).

3.2.3 The Mixed-effects Model

The last model we introduce is a mixed-effects model, that integrates the information from the K unsmoothed fMRI images, by postulating that these K unsmoothed 3-dimensional Z-statistic datasets are originated from one underlying true, smooth Z-statistic image. In what follows, we will refer to this model as the Mixed model. The major difference between this model and the models previously described is that this model postulates that there exists only one smoothed 3-dimensional Z-statistic image, of which the K unsmoothed fMRI images are K noisy versions. In contrast, the MCAR model and the three proposed spatially adaptive MCAR models all assume and yield inference on K smoothed 3-dimensional Z-statistic images. The specification of the Mixed model is as follows:

$$\begin{aligned} Y_{ik} &= \mu_i + \epsilon_{ik} \qquad \epsilon_{ik} \overset{iid}{\sim} \mathrm{N}(0, \tau^2) \\ &= \beta_0 + b_{0i} + \epsilon_{ik},\ i = 1, ...N,\ k = 1, 2, \end{aligned}$$

where $\boldsymbol{\mu} = (\mu_1, ..., \mu_N)^T$ is the resulting smoothed image, b_{0i} accounts for the between-voxel variance and τ^2 for the within-voxel variance. The spatial dependence between the voxels is captured through the conditional autoregressive (CAR) model, which we use as a specification for the prior on $\boldsymbol{b}_0 = (b_{01}, ..., b_{0N})^T$. Specifically,

$$\left[b_{0i} \mid \boldsymbol{b}_{(-0i)}, \sigma^2\right] \sim \mathrm{N}\left(\sum\nolimits_j \frac{w_{ij}}{w_{i+}} b_{0j}, \frac{\sigma^2}{w_{i+}}\right). \tag{3.12}$$

Further, we conclude the Mixed model specification by placing priors on β_0 and τ^2: $\pi(\beta_0) \propto 1$ and $\tau^2 \sim \mathrm{IG}(1, 1)$.

3.3 Simulation Studies

In this section, we describe two simulation studies to evaluate the performance of these six models. The first study assumes independent image pairs while the second assumes correlated pairs. The six models include: the CWAS model applied to each image independently, the MCAR model, the three different types of spatially adaptive MCAR models, and the Mixed model. Specifically, in the first simulation study, fMRI data sets are simulated independently from an R package with known activation regions. In the second simulation study, 2-dimensional images are simulated from a Gaussian Markov Random Field (GMRF) with known correlations for each image pair. In the first simulation study, we compare the models' performance in terms of the false negative rate (FNR) and the false positive rate (FPR). In the second study, the comparison is based on the average mean squared error (MSE) and bias of the mean intensities and correlations. We choose different performance measures since 1) the activated voxels are pre-determined in the first study but not in the second, and 2) the underlying true values of the mean intensities and correlations are known in the second study but not in the first.

3.3.1 Simulation Study 1

In this first experiment, we simulate 4-dimensional fMRI time series data sets using the same simulation strategy and settings used in Chapter II. Specifically, we employ the R neuRosim package (Welvaert et al., 2011) and we refer the reader to Section 2.4 for more details on the simulation settings. We consider signal to noise ratios (SNR) of 1, 2, and 3. For each SNR, we simulate 50 independent fMRI data sets and for each simulated 4-dimensional fMRI data set, we generate a 3-dimensional Z-statistic image using FEAT in the FSL (Smith et al., 2004) software package. As our primary goal is to smooth the data, we turn off the spatial smoothing in the preprocessing stage, and thus the generated 3-dimensional Z-statistic images are unsmoothed. We apply all

six models (CWAS, MCAR, SAMCAR I, SAMCAR II, SAMCAR III and the Mixed model) to two randomly chosen unsmoothed Z-statistic images from the 50 simulated images for each SNR. We repeat this procedure 50 times. All parameter estimates are obtained within the Bayesian framework, and posterior inference is based on 10,000 MCMC samples: each algorithm is run for 25,000 iterations, with a burn-in of 10,000 iterations.

We use the Bayesian decision rule proposed in Chapter II Section 2.2.3 to classify the voxels into three classes: activated voxels, deactivated voxels and null voxels. This decision rule is designed for pre-surgical fMRI analysis, as it allows for control of both false positives and false negatives. Using the underlying true class of the voxel and the estimated classes of the voxel determined with our Bayesian decision rule, we can further label each voxel as true positive (TP), true negative (TN), false positive (FP) or false negative (FN). A voxel is labeled as a FP if it is truly null but falsely classified as activated or deactivated, while it is labeled as a FN if it is truly activated or deactivated but falsely classified as null. If its true class and its estimated class agree, then an activated or deactivated voxel is labeled as TP while a null voxel is labeled as TN.

In this simulation study, we evaluate the performance of all 6 models based on their false negative rate (FNR) and false positive rate (FPR):

$$\begin{aligned} \text{FNR} &= \text{FN}/(\text{FN}+\text{TP}), \\ \text{FPR} &= \text{FP}/(\text{FP}+\text{TN}). \end{aligned}$$

For each model, except the Mixed model, the FNR is obtained by averaging the model's FNR across the two images and the 50 simulations. As the Mixed model only yields one smoothed image per pair of images, its FNR is the average of the FNRs across the 50 simulations.

Figure 3.1 presents, for each SNR setting, the FNR and FPR of the proposed SAMCAR models against the CWAS model and the standard MCAR model when the threshold used to classify voxels into null and non-null varies from 0 to 1. In general, the MCAR and Mixed models outperform the other models across all SNRs. The CWAS model is slightly better than the MCAR and Mixed model if we only focus on the FNR for high SNRs and low threshold values. Among all proposed SAMCAR models, SAMCAR III has a relatively better performance overall, while SAMCAR I achieves similar performance to the MCAR model in terms of FNR when the SNR is set to 3 at small threshold values.

3.3.2 Simulation Study 2

To investigate the performance of the six models when the multiple images are correlated, we simulate pairs of correlated images with three chosen correlations: 0.25, 0.5, 0.75. To simulate correlated images we proceed as follows. First, we simulate two independent 2-dimensional 50×50 images, $\boldsymbol{\phi}_1$ and $\boldsymbol{\phi}_2$, using the joint distribution implied by the proper CAR model:

$$\boldsymbol{\phi}_i \sim \text{MVN}\left(0, \theta^2[(\boldsymbol{D} - \alpha \boldsymbol{W})]^{-1}\right), \ i = 1, 2,$$

where $\theta^2 = 4$, $\alpha = 0.999$. We then construct the pair of correlated images by setting: $\boldsymbol{\mu}_1 = \boldsymbol{\phi}_1$ and $\boldsymbol{\mu}_2 = t\boldsymbol{\phi}_1 + (1-t)\boldsymbol{\phi}_2$. For different choices of t, we obtain different correlations between $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$. Table 3.1 shows the expected correlations and the observed correlations between the simulated images $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$ for different choices of t. Finally, after having simulated the smoothed images, we generate the unsmoothed images by setting Y_{ij} as: $Y_{ij} \sim \text{N}(\mu_{ij}, \tau^2), \ i = 1, ..., 2500, \ j = 1, 2$ where $\tau^2 = 0.5$ or 1. Figure 3.2 displays both the simulated independent images ($\boldsymbol{\phi}_1$ and $\boldsymbol{\phi}_2$), and the simulated correlated images (one common $\boldsymbol{\mu}_1$ and three different $\boldsymbol{\mu}_2$s under different

choices of correlations).

We compare the various models in terms of average mean squared error ($\overline{\text{MSE}}$) and average absolute bias ($\overline{|\text{Bias}|}$) of the estimates of the μ_{ij}s, and for the correlation between $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$, respectively. To compute these two summary statistics, we use the following definitions:

$$\begin{aligned}
\overline{\text{MSE}} &= (NKM)^{-1} \sum_{i=1}^{N} \sum_{k=1}^{K} \sum_{s=1}^{M} (\hat{\mu_{ik}}^s - \mu_{ik})^2, \\
\overline{|\text{Bias}|} &= (NKM)^{-1} \sum_{i=1}^{N} \sum_{k=1}^{K} \sum_{s=1}^{M} |\hat{\mu_{ik}}^s - \mu_{ik}|,
\end{aligned}$$

where $\hat{\mu_{ik}}^s$ is the posterior mean of μ_{ik} for voxel i in the kth image and sth simulation.

Figure 3.3 shows the average MSE and average absolute bias of the estimates of the μ_{ik}s for all six models for simulations generated under three different correlations (0.25, 0.5 and 0.75) and two different error variances τ^2s (0.5 and 1). In terms of estimation accuracy for μ_{ij}s, we can see that the MCAR model performs better than all the other models under all scenarios. Next in terms of performance, are the CWAS model and the SAMCAR I model which yield average MSEs and average absolute biases close in value to those of the MCAR model. The mixed-effects model yields the highest average MSE and absolute bias. We believe that this is due to the fact that the mixed-effects model assumes that the unsmoothed images originate from one smoothed image while the simulated data instead arises from two different smoothed images.

The average MSE and bias for the correlation between the simulated $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$ are presented in Figure 3.4. Overall, the MCAR model performs better than the other models. We note that the extent of random noise added to each $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$, which is controlled by τ^2, seems to play an important role on the performance of these models. While the performance of the CWAS model is stable under different settings of correlations and error variances when we focus on the FNR and the estimation

accuracy for μ_{ij}s, we see that the estimation accuracy of the CWAS model falls off rapidly as both the noise and correlation increase. In general, as the amount of noise and correlation increase, the performance of all models decreases, except for the SAMCAR III model. It is interesting to note that the SAMCAR III model outperforms the other five models in terms of estimation accuracy of the correlation under high correlation (correlation=0.75).

3.4 Application

In this section, we present the results obtained when we apply the three proposed spatially adaptive SAMCAR models and the Mixed model to the fMRI datasets from two patients. For completeness in the comparison, we also apply the CWAS and the MCAR model to the patients images. For each patient, we have two Z-statistic fMRI images obtained via mass univariate GLM analysis without spatial smoothing. Both patients' fMRI data sets have the same dimension: $64 \times 64 \times 40$. Figure 3.5 shows two saggital slices of the unsmoothed Z-statistic images for each patient. It should be noted that the fMRI datasets for the first patient are acquired under different experimental paradigms, i.e., tongue twister and sentence completion, whereas the fMRI datasets for the second patient are acquired under the same experiment. For the first subject, there is little empirical correlation between the two images (overall correlation = -0.05). For the second subject, the empirical correlation is a bit larger (overall correlation = 0.35). We also anticipate that different activated and deactivated regions will be detected by the two fMRI images acquired under the two different paradigms for the first patient, while the two fMRI images acquired under the same paradigm for the second patient would identify similar activated and deactivated regions.

We apply all six models to the pairs of Z-statistic fMRI images from both patients, while we fit the CWAS model to each Z-statistic image separately. To classify voxels

into activated and deactivated voxels, we use the loss function described in Section 2.2.3 with a threshold set to 0.2, corresponding to a penalty on false negatives that is 11 times larger than for the false positives. In short, a voxel is classified as activated when its estimated mean intensity is greater than zero and the normalized and standardized mean intensity m_i satisfies $m_i/\max(\mathbf{m}) \geq 0.2$, while it is deemed deactivated when its estimated mean intensity is less than zero and $m_i/\max(\mathbf{m}) \geq 0.2$, where $\mathbf{m} = (m_1, m_2, ...m_N)^T$.

Figure 3.6 shows the smoothed images, i.e., $\widehat{\boldsymbol{\mu}_1}$ and $\widehat{\boldsymbol{\mu}_2}$, estimated using the six models. Figure 3.6 reveals that the smoothing effect of these models is different, with MCAR and SAMCAR I smoothing the least, CWAS and SAMCAR II slightly more, and SAMCAR III and Mixed the most. Further, Figures 3.7 and 3.8 present results on the loss function scale with color overlay for the activated and deactivated regions for patient one and two, respectively. These two figures clearly show that different activated and deactivated regions are identified by the two images relative to the first patients while similar activated and deactivated regions are identified by the two images from the second patient, as we have anticipated given the paradigm used in obtaining these images. While ground truth on the classification results obtained is not available, the activated and deactivated regions for patient one under the tongue twister paradigm identified using CWAS (Image 1 in the first row of Figure 3.7) have been previously confirmed by intra-operative ESM. Thus, we consider these regions for patient one's first image as a positive control and we compare the regions identified by the other models with the former.

SAMCAR II gives results most similar to CWAS. Recall that SAMCAR II is a multivariate extension of the CWAS model, and the difference between SAMCAR II and CWAS is that the former introduces voxel-specific correlation parameters in its conditional covariance matrix. Results from both patients indicate that these additional parameters do not improve the estimation of the mean intensities. One

potential explanation for this is the large number of free parameters needed to be estimated for this model and the relatively small number of observations. The results from MCAR and its extension, SAMCAR I, are also comparable, and they both identify fewer and smaller activated and deactivated regions than SAMCAR II and CWAS. Similarly, this result indicates that the additional parameters introduced in the SAMCAR I model do not significantly translate into an improvement in inference for the mean intensity. An interesting finding is obtained when we compare the SAMCAR III model with the Mixed model. For the first patient where the correlation between the two images is low, the smoothing effect of SAMCAR III resembles that of the CWAS model and is different from that of the Mixed model. On the other hand, for the second patient, with larger between-image correlation, the SAMCAR III model and the Mixed models yield similar smoothing results. This finding suggests that the SAMCAR III model provides an amount of smoothing that is adaptable to various between-image correlation levels. Lastly, we note that since the Mixed model assumes that the pair of images for each patient have the same expected activated and deactivated regions, it always produces the same inferences for the pair of images.

3.5 Discussion

In this chapter, we have proposed different smoothing approaches to integrate information across multiple fMRI images exploiting the potential correlation among images. These models have all been formulated with the goal of improving statistical inference on activated and deactivated regions by potentially leveraging local correlations within and across images from the same patient. Our first sets of model was developed on the assumption that correlation among images might vary spatially. Building upon this assumption, we propose three spatially adaptive multivariate conditional autoregressive (SAMCAR) models, which can be considered as extensions of the MCAR model (Carlin and Banerjee, 2003; Gelfand and Vounatsou, 2003), the

CWAS model (Liu et al., 2016), and the RH model (Reich and Hodges, 2008), respectively. We also proposed a mixed-effects model that assumes that the multiple fMRI images are all noisy versions of one underlying smooth fMRI image.

Compared with the CWAS and MCAR models, the proposed spatially adaptive multivariate conditional autoregressive models, named SAMCAR I, II and III, have more degrees of freedom, allowing the correlation between fMRI images to vary in space. Applying these models to simulated and real data, however, revealed several challenges. Specifically, we found that the additional spatially varying parameters introduced to SAMCAR I, II and III do not significantly lead to an improvement in the estimation of the mean intensity. This could be due to an insufficient number of observations, an excessive model flexibility, or the pseudo-likelihood approximation. We also note that a better assessment of the proposed models would require further electrical stimulation mapping (ESM) validations.

One interesting finding is that the SAMCAR III model does adaptively recover the underlying correlation embedded in the images. This could be attributed to the imposed symmetric feature on its joint covariance matrix. Such a structured joint covariance matrix makes it straightforward to understand since it directly models the covariance matrix between multiple fMRI datasets. On the other hand, the conditional covariance matrices in both SAMCAR I and II cannot be directly interpreted as the covariance matrix across multiple fMRI datasets.

t	$\frac{1}{\sqrt{15}+1}$	$\frac{1}{\sqrt{3}+1}$	$\frac{3}{\sqrt{7}+3}$
Expected Correlation	0.25	0.5	0.75
Observed Correlation	0.29	0.51	0.74

Table 3.1: Expected correlation and observed correlation between the simulated images $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$ with different choices of t

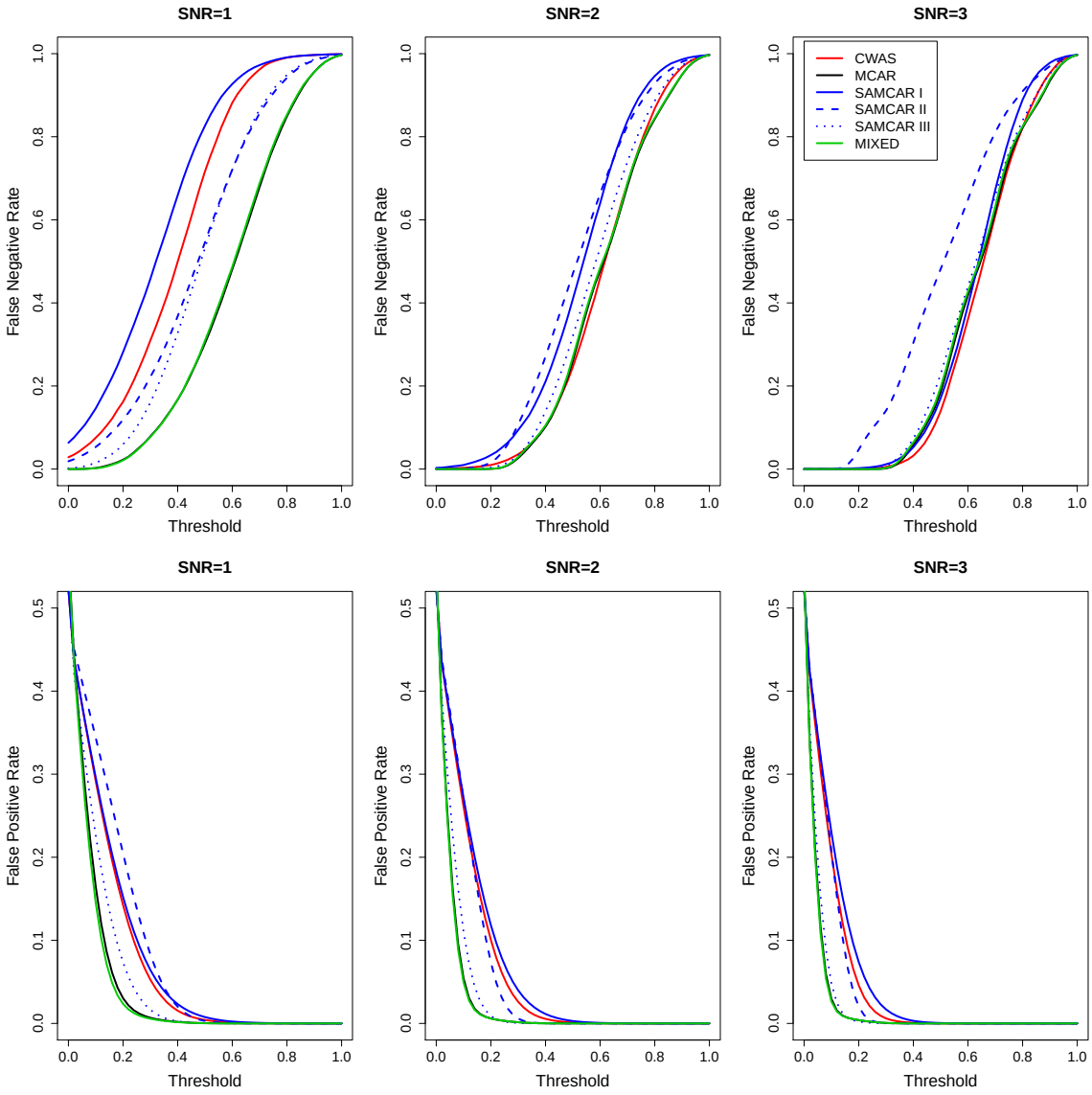

Figure 3.1: False negative and false positive rates vs. loss function threshold for simulation study 1. For each signal to noise ratio, the false negative rate and false positive rate are averaged over two fMRI data sets and 50 simulations, except for the Mixed model that is averaged over the 50 simulations.

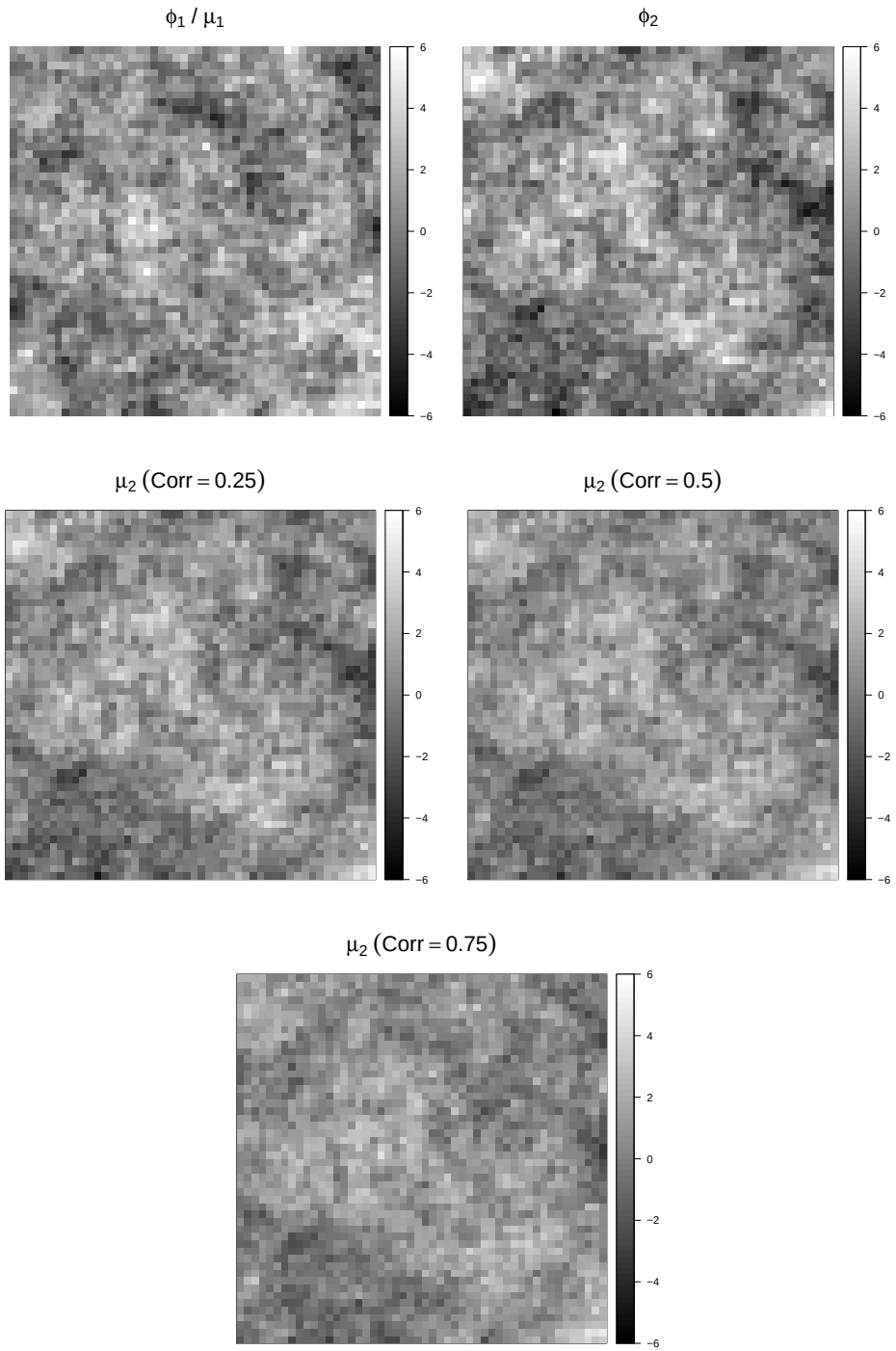

Figure 3.2: Simulated images in simulation study 2. The first row is the simulated $\boldsymbol{\phi}_1$ / $\boldsymbol{\mu}_1$ and $\boldsymbol{\phi}_2$. The second and third rows are the simulated $\boldsymbol{\mu}_2$ at different chosen correlations.

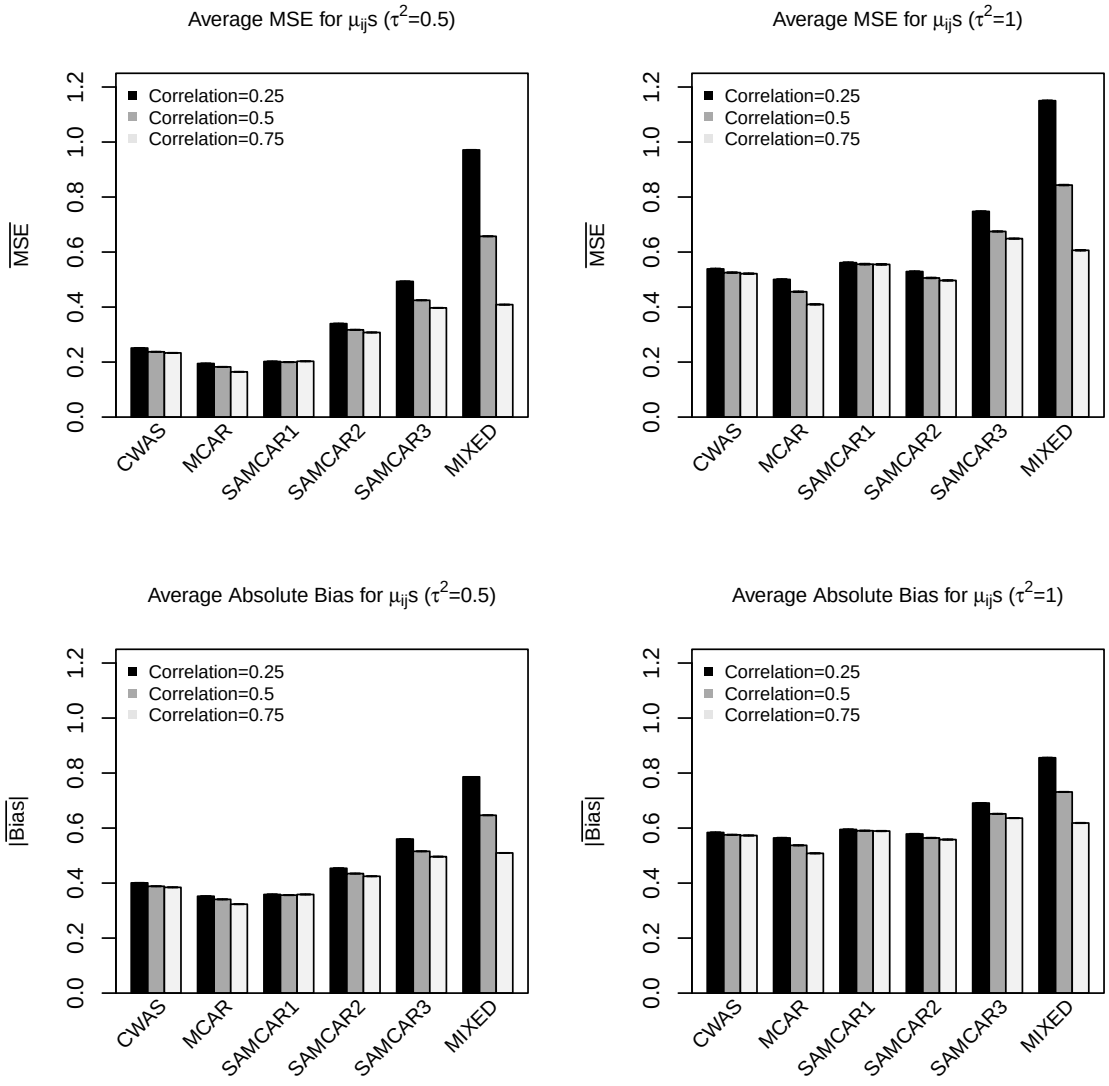

Figure 3.3: Average MSE and average absolute bias for posterior estimates of μ_{ik}s from all six models under three different correlations (0.25, 0.5 and 0.75) and two different τ^2s (0.5 and 1) in simulation study 2. Standard error bars are included (too small to detect).

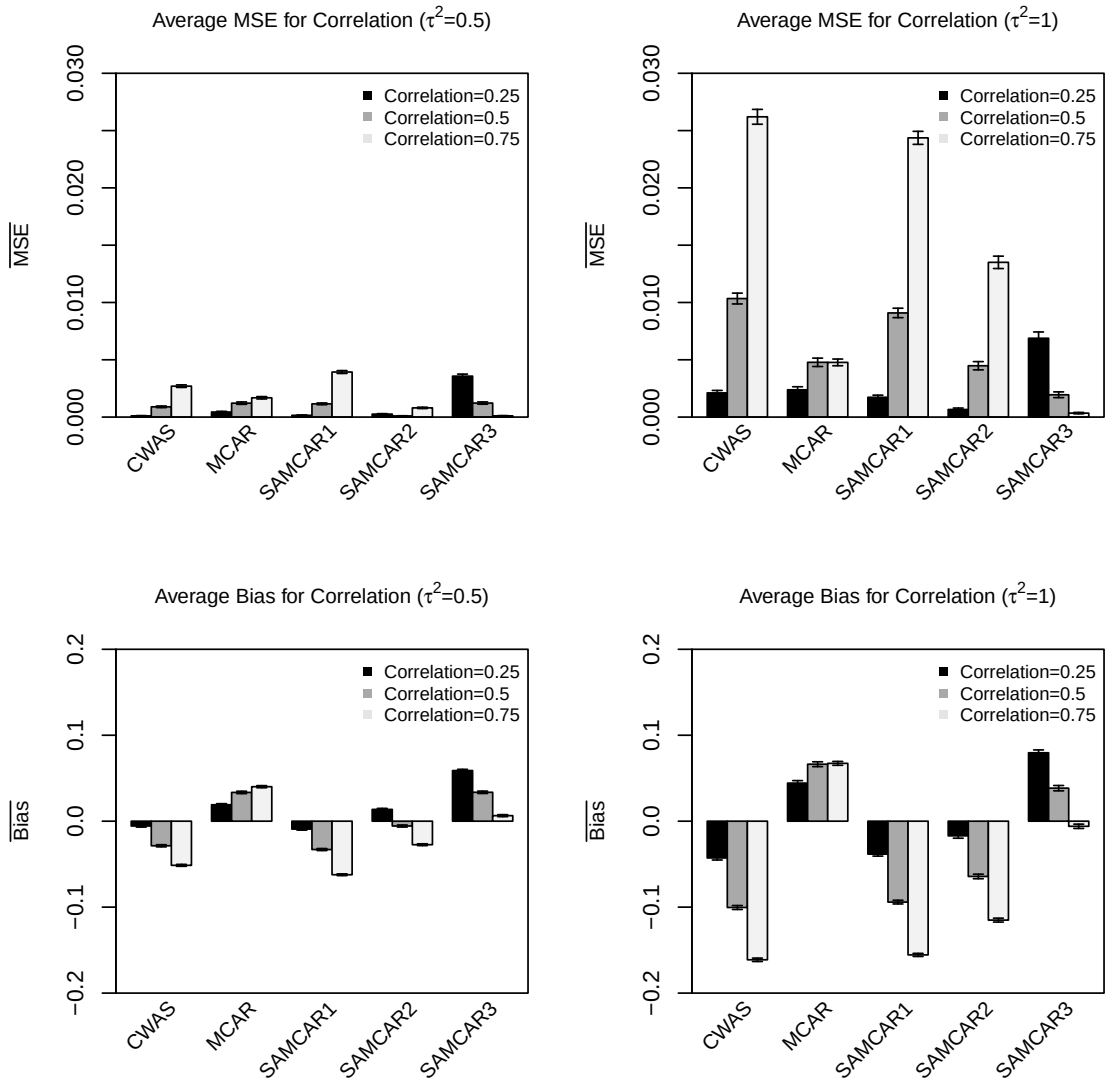

Figure 3.4: Average MSE and average bias for the correlation between the posterior estimates of $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$ from all six models under three different correlations (0.25, 0.5 and 0.75) and two different τ^2s (0.5 and 1) in simulation study 2. Standard error bars are included.

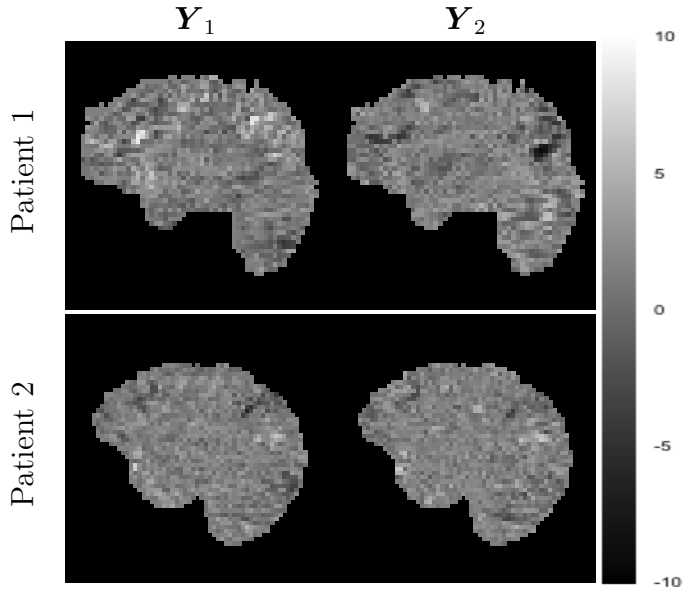

Figure 3.5: Patients' fMRI data. Top row: Patient 1's two fMRI Z-statistic images acquired at two different experimental paradigms. Bottom row: Patient 2's two fMRI Z-statistic images acquired at the same experimental paradigm.

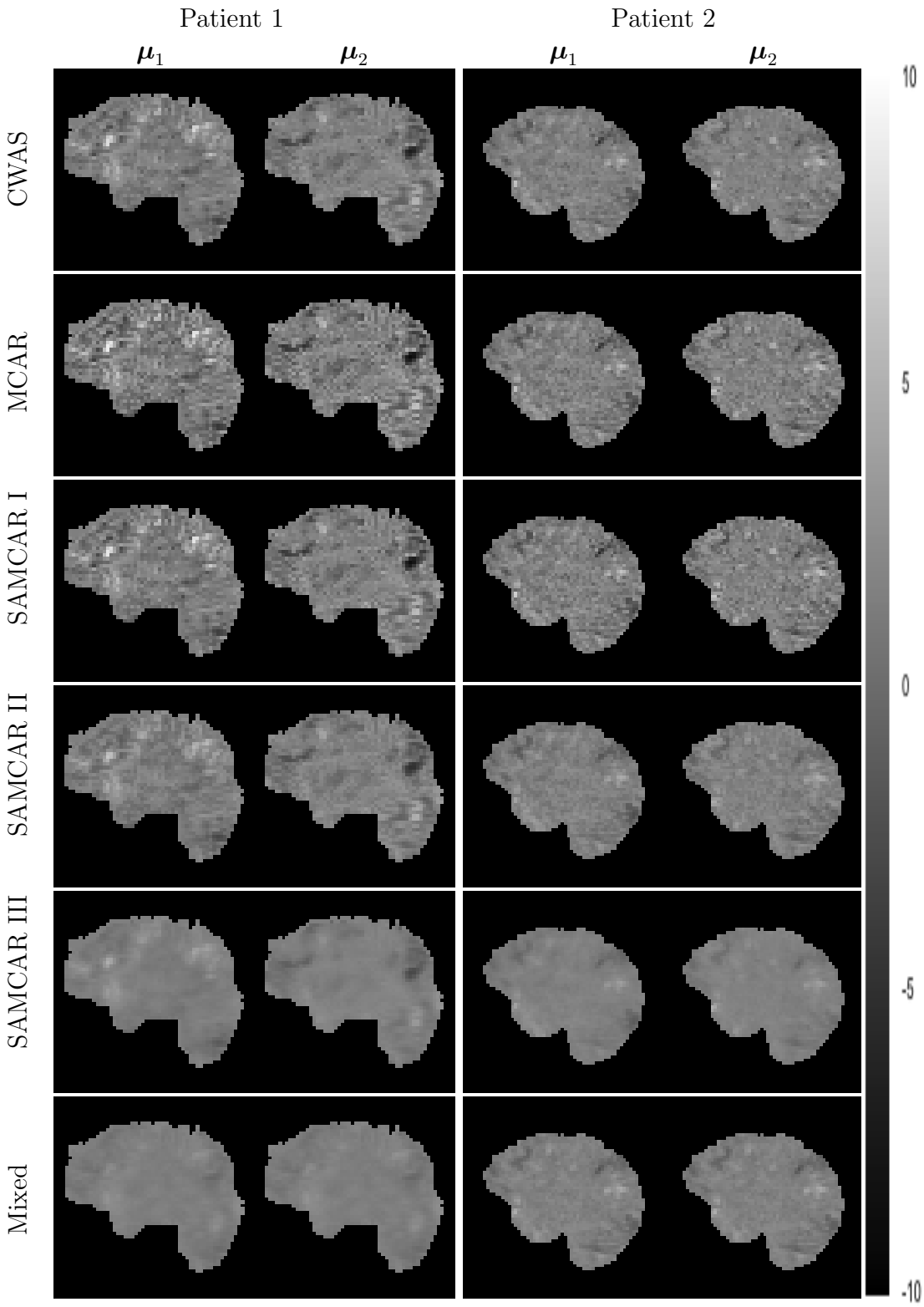

Figure 3.6: Results from the six models for the two patients. The first two columns shows the smoothed images, i.e., $\widehat{\boldsymbol{\mu}_1}$ and $\widehat{\boldsymbol{\mu}_2}$ for patient 1. The third and fourth columns shows the smoothed images, i.e., $\widehat{\boldsymbol{\mu}_1}$ and $\widehat{\boldsymbol{\mu}_2}$ for patient 1. The name of each row indicates the model applied.

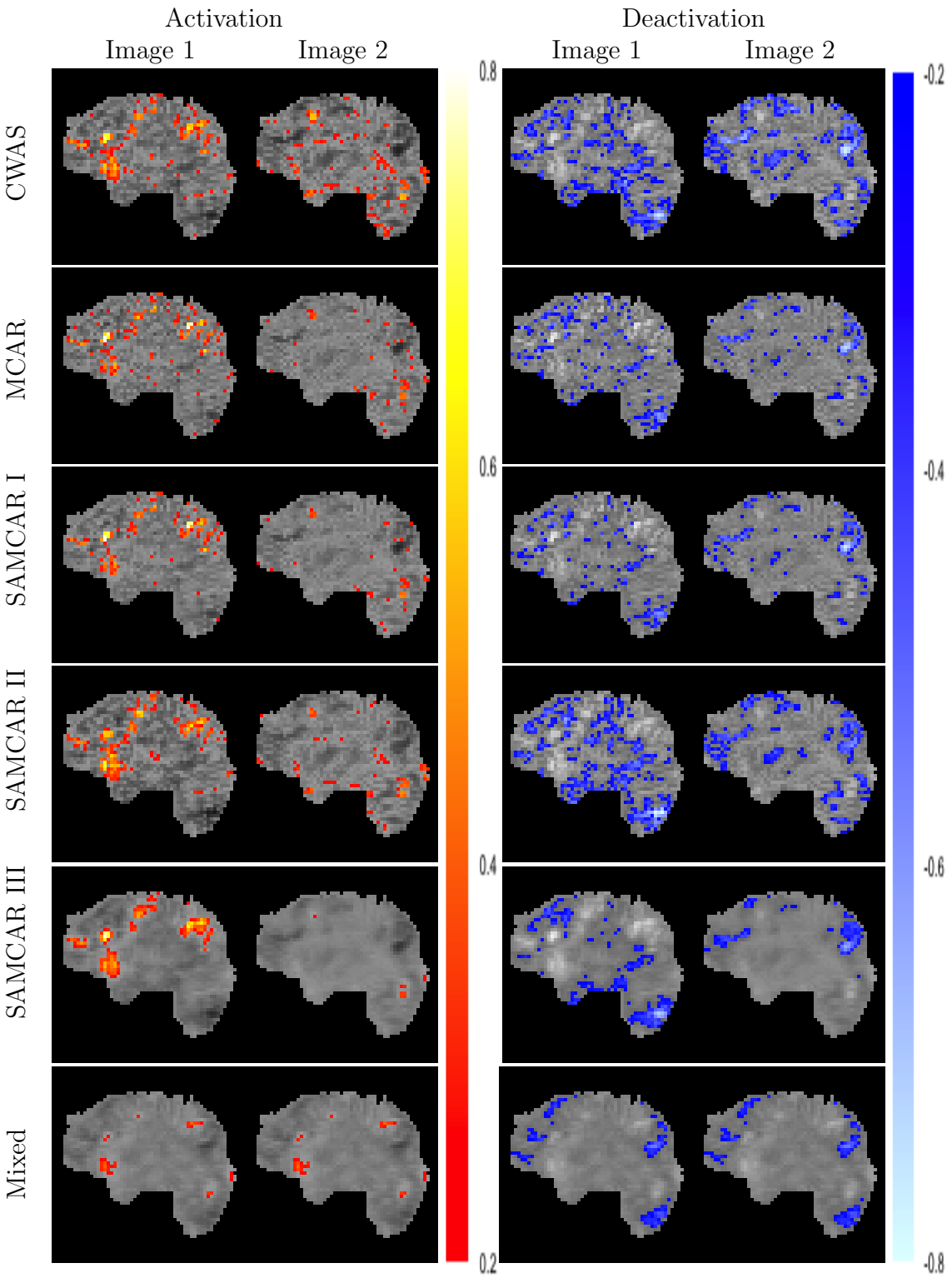

Figure 3.7: Patient 1's results on the loss function scale. Voxels with loss function scale above 0.2 (below -0.2) are colored as activated (deactivated) regions. Values above 0.8 (below -0.8) are mapped to 0.8 (-0.8) to give a better dynamic range of colors. Red to yellow (Dark blue to bright blue) denote increasingly strong activation (deactivation).

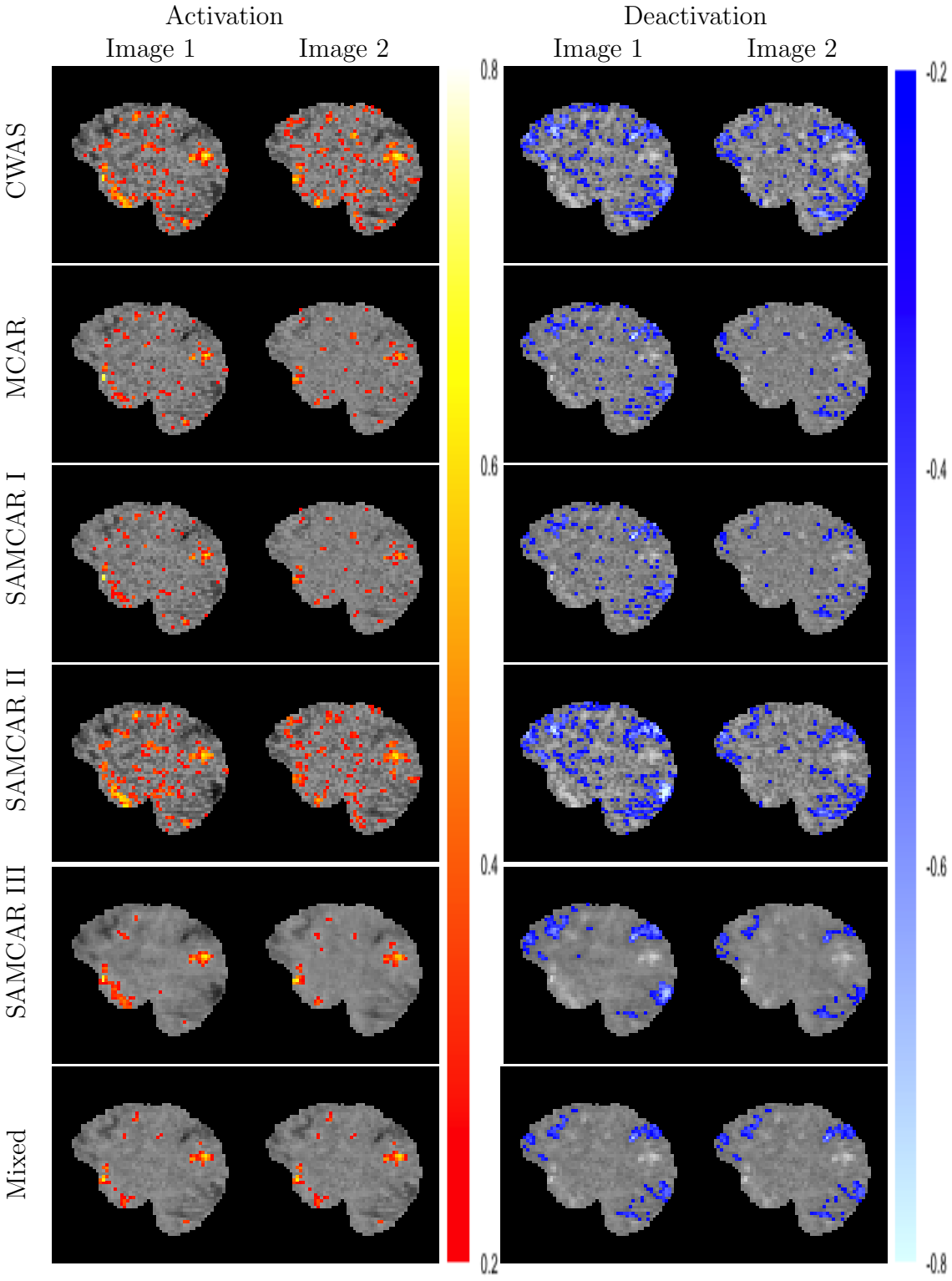

Figure 3.8: Patient 2's results from the six models on the loss function scale with color overlay for activated and deactivated regions (see Figure 3.7 for details).

CHAPTER IV

A Mixed-Effects Model with Spatially Varying Coefficients with Application to Multi-resolution FMRI Data

4.1 Introduction

One of the goals during neurosurgery, brain tumor resections in particular, is to accurately resect as much of an intra-axial brain lesion as possible while minimizing damage to surrounding healthy tissue (Tharin and Golby, 2007). To achieve this goal, precise information about a patient's specific structural and functional anatomy is essential. Although neurosurgeons use direct electrical stimulation mapping (ESM) during surgery, there is growing interest in using less invasive methods for pre-surgical planning and intra-operative guidance (Hirsch et al., 2000; Sunaert, 2006).

One such non-invasive method is functional Magnetic Resonance Imaging (fMRI) that is widely used in neuroscientific studies and is increasingly being proposed for pre-surgical planning (Kekhia et al., 2011). However, it has not been widely adopted. Pre-surgical fMRI mapping of functional brain regions may help surgeons in different ways: 1) optimize the surgical strategy (Håberg et al., 2004; Tieleman et al., 2009); 2) guide the deployment of intra-operative ESM; and 3) predict potential deficits in neurological and cognitive functions due to surgery or continued tumor growth.

The goals of pre-surgical fMRI analysis are fundamentally different from those of neuroscientific fMRI analysis. Furthermore, there is a large body of research on neuroscientific fMRI data analysis; however, little literature exists for pre-surgical fMRI data analysis, especially from a statistical point of view. To date, there are no guidelines available for pre-surgical fMRI analysis (Stippich, 2007a) and broad adoption of fMRI into standard clinical practice is precluded by a relative lack of clinical studies and trials proving definitive benefits for pre-surgical planning and intra-operative neuro-navigation (Sunaert, 2006). When analyzing pre-surgical fMRI data, one should always be aware of the applicability and limitations of neuroscientific fMRI analyses (Haller and Bartsch, 2009; Bartsch et al., 2006). In particular, unlike neuroscientific analysis where inferences are made on data from groups of subjects (Stippich, 2007b), pre-surgical fMRI measurements and analyses are performed on individual patients (Stippich, 2007a), leading to a demand for more sophisticated models to compensate for the (relative) lack of data. Furthermore, while spatial smoothing using a Gaussian kernel is a common preprocessing step for neuroscientific fMRI analysis, it may not be appropriate or desirable for pre-surgical fMRI analysis. In fact, whereas this preprocessing step increases the signal-to-noise ratio (SNR) for activations greater than or equal to the size of the smoothing kernel (FWHM) and guarantees that the assumptions of Gaussian random field theory hold (Worsley et al., 1996), smoothing blurs boundaries between functionally activated/deactivated regions and null regions of the brain and may cause small activated/deactivated regions to shrink or even disappear. As a result, truly activated or deactivated regions might be deemed as neither activated nor deactivated—leading to false negatives—which is a concern for pre-surgical planning where spatial accuracy of the functionally eloquent regions of the brain is critical. Lastly, while neuroscientific fMRI analysis is designed to prevent false positives in the hypothesis testing framework, false negatives are of equal or greater concern in pre-surgical planning (Bartsch et al., 2006; Haller

and Bartsch, 2009; Durnez et al., 2013). Incorrectly classifying functionally eloquent brain regions as irrelevant may lead to irreversible brain damage during surgery.

The spatial precision with which activated and deactivated regions can be detected depends largely on the spatial resolution of the fMRI scan. The spatial resolution that can reasonably be achieved within an fMRI image is determined by many study parameters, such as the main magnetic field strength, volume acquisition time, and volume coverage (Yoo et al., 2004). It is measured by the size of voxels, which are three-dimensional rectangular prisms. The X, Y, Z dimensions of voxels are determined by in-plane resolution, slice thickness and inter-slice gap and can be different in the three directions. Increasing the spatial resolution of fMRI is, at least in theory, relevant for clinical applications to improve spatial accuracy. For example, greater spatial fMRI resolution may improve the co-registration of low-resolution functional images with high-resolution anatomical images used for intra-operative neuro-navigation. In fact, Yoo et al. (2004) investigated the importance of spatial resolution of fMRI and its relation with the precision of functional localization for surgical planning. They concluded that fMRI data acquired at 2-mm isotropic resolution, which is higher than the resolutions used in typical fMRI studies, significantly enhances the spatial precision of activation patterns. However, greater spatial resolution comes at a cost: a loss in the SNR, due to both a loss in the measured signal, which decreases approximately linearly with the voxel size, and an increased relative contribution of thermal noise that increases non-linearly at higher spatial resolutions. This is also why high-resolution task-based fMRI data, recorded by simultaneous multi-slice (SMS) imaging for example, tends to require an equal amount of smoothing as low-resolution recordings to achieve comparable results (Harms et al., 2013). In addition, improved spatial resolution of fMRI data brings new challenges with respect to statistical analysis as it requires appropriate approaches to best take advantage of the richer data.

In this paper, we aim to develop a statistical approach for pre-surgical fMRI analysis when the fMRI data available for a brain tumor patient is obtained at two different spatial resolutions: a commonly used spatial resolution for conventional studies and a higher spatial resolution recorded through simultaneous multi-slice (SMS) accelerated imaging (Feinberg et al., 2010) which can be used not only to accelerate fMRI acquisitions temporally but also to record fMRI data at higher spatial resolution. The example data set used in this manuscript is from an fMRI study where the patient performed the same task during two separate sessions during which the fMRI data was recorded at two resolutions: the standard resolution with a total number of $64 \times 64 \times 48$ voxels, each voxel having dimensions $3 \times 3 \times 3.45$ mm^3; and at a larger resolution with a total of $120 \times 120 \times 62$ voxels, with voxel dimensions $1.8 \times 1.8 \times 2.30$ mm^3.

Liu et al. (2016) proposed a Bayesian model for pre-surgical fMRI analysis that employed a novel spatially adaptive conditionally autoregressive (CAR) model for smoothing, where the amount of smoothing varies spatially and is driven by the data, thus reducing the blurring of boundaries between (de)activated regions and null regions. Subsequently, they adopted a Bayesian theoretical decision approach to control both the false negative rate (FNR) and the false positive rate (FPR) in an asymmetric manner. However, their model does not jointly model both high and standard resolution fMRI data.

Inferring upon a spatial random field using data collected at different spatial resolutions is a problem often encountered in spatial statistics and is referred to as the "change of support" problem (COSP) (Gelfand et al., 2001). A variety of approaches have been proposed to address this problem (see Gotway and Young (2002) for an overview). Point Kriging and block Kriging are the commonly used solutions to the point-to-point COSP and the point-to-area COSP, respectively (Banerjee et al., 2014). Addressing a more general class of COSP problems, Gelfand et al. (2001) proposed

a unifying Bayesian hierarchical approach for prediction from point-to-point, point-to-area, area-to-point, and area-to-area, while Wikle and Berliner (2005) presented a Bayesian hierarchical method to combine information across different scales, assuming the existence of a true latent process. More recently, with an application to air pollution, Berrocal et al. (2010) proposed a model-based strategy to downscale the data from areal level to point level, using a linear regression model with spatially varying coefficients (Gelfand et al., 2003). The ultimate goal of most of these approaches, which were developed for continuous spatial processes, typically modeled as spatial Gaussian processes, was prediction of the spatial process at unsampled locations. As brain images are registered to a common anatomical reference space, leaving no unobserved regions, there is no need for prediction. Thus, the aforementioned approaches for handling the COSP are not necessary for neuroimaging data. In addition, as fMRI data is specified at the voxel-level for a very large number of voxels, modeling this spatial data using a continuous spatial process might not be appropriate, let alone be computationally tractable, as the likelihood evaluation would require computation of the inverse and determinant of a very large covariance matrix.

To address these issues and incorporate the information from two fMRI brain images acquired at different spatial resolutions, we develop a mixed-effects model with spatially varying coefficients (SVC), that is closely related to that of Berrocal et al. (2010). Specifically, prior to model implementation, we register the motion- and distortion-corrected low-resolution functional image space to the corresponding fMRI data recorded at higher spatial resolution. Thus, for each voxel in the high-resolution image, there is a corresponding voxel in the low-resolution image (the low resolution image is interpolated into the space of the high resolution image using nearest-neighbor interpolation). We will refer to this interpolated low resolution image as the nearest-neighbor-interpolated low-resolution image. Following Berrocal et al. (2010), we regress the high-resolution image on the nearest-neighbor-

interpolated low-resolution image via a Bayesian hierarchical model with spatially varying coefficients. However, in our model only the slope terms vary spatially. We account for their spatial structure by placing a conditional autoregressive (CAR) prior (Besag et al., 1991) on them. This model specification, besides taking into account the nature of the neuroimaging data, allows for a computationally efficient algorithm to be developed.

The remainder of the article is organized as follows. In Section 4.2, we describe our mixed-effects model with spatially varying coefficients that integrates fMRI data with different resolutions. We briefly explain how we identify activated and deactivated regions using a Bayesian decision approach originally proposed by Liu et al. (2016). In Section 4.3, we conduct a simulation study to compare the performance of our proposed SVC model with that of the model by Liu et al. (2016). In Section 4.4, we apply our proposed method to a patient's multi-resolution pre-surgical fMRI data set. Finally, conclusions and future directions are given in Section 4.5.

4.2 Methods

In this section, we first describe how we register the low- and high-resolution data and then we propose a model that incorporates the data at the two resolutions.

4.2.1 Registration

In order to incorporate data at the two resolutions into the model, we first register the low-resolution Z-statistic image to the high-resolution Z-statistic image, thus creating an interpolated low-resolution Z-statistic image that has the same size and shape as the high-resolution Z-statistic image. This is accomplished by creating an unbiased within-subject template as the common space for the low and high resolution functional data (Reuter et al., 2010; Reuter and Fischl, 2011; Reuter et al., 2012), using the motion- and distortion-corrected functional images, inverse consistent, rigid

body (6 degrees of freedom) image registration and nearest neighbor interpolation. When considering a 3-dimensional image, interpolation refers to the process used to transform a discrete image defined at a smaller set of coordinate locations to a larger set of coordinate locations (Parker et al., 1983). We choose the nearest neighbor interpolation method to avoid smoothing which is, instead, integrated in our model. Any interpolation method other than nearest neighbor interpolation, such as trilinear interpolation and spline interpolation, inherently smoothes the data. The unbiased within-subject template, generated by mri_robust_template, part of FreeSurfer (http://surfer.nmr.mgh.harvard.edu/; Fischl (2012)), is used as the undistorted reference space for the pre-processing and statistical analysis of the fMRI data acquired at both resolutions. Denoised Z-statistic images from the two resolutions are created by first removing motion components in the preprocessing and then applying the mass univariate GLM approach in the statistical analysis. Preprocessing and mass univariate GLM analysis is performed using FEAT, part of fMRIB Software Library (FSL) (http://www.fmrib.ox.ac.uk/fsl/, Smith et al. (2004); Woolrich et al. (2009)). As the result of the interpolation, each voxel in the interpolated low-resolution image corresponds to a unique voxel in the high-resolution image.

4.2.2 The Proposed Model

Consider two unsmoothed 3-dimensional Z-statistic images, one being the high-resolution image and the other one being the interpolated low-resolution image. Let X_i and Y_i indicate the Z-statistic value for voxel i in the interpolated low-resolution image and the high-resolution image, respectively, $i = 1, ..., N$ where N is the total number of voxels. Let $\boldsymbol{X} = (X_1, ..., X_N)^T$ and $\boldsymbol{Y} = (Y_1, ..., Y_N)^T$ represent the interpolated low-resolution image and the high-resolution image, respectively. We assume that the Z-statistic values in the high-resolution image follow, independently, a normal distribution, i.e., $Y_i \sim N(\mu_i, \tau^2)$ or $Y_i = \mu_i + \epsilon_i$ with $\epsilon_i \overset{iid}{\sim} \mathrm{N}(0, \tau^2)$ for $i = 1, ..., N$.

In the formulation $\boldsymbol{Y} = \boldsymbol{\mu} + \boldsymbol{\epsilon}$, $\boldsymbol{\mu} = (\mu_1, ..., \mu_N)^T$ represents the smoothed Z-statistic image.

We relate $\boldsymbol{\mu}$ to $\boldsymbol{X}$ by assuming that at each voxel i, $\mu_i = \beta_0 + (\beta_1 + b_{1i})X_i$. The spatial association between the high-resolution image and the low-resolution image is believed to be different at each voxel, thus voxel specific slopes, b_{1i}s, are introduced to capture the spatially varying associations. We note that a model formulation with both spatially varying intercepts and slopes leads to insufficient smoothing of the image, as we have observed empirically. To be more specific, the low resolution data can only capture the large scale spatial structure of the high resolution image data, and when a spatially-varying intercept term is introduced, that term tends to capture the fine scale spatial structure of the high resolution image, leading to almost no smoothing, and may also cause the spatially-varying slopes to be less variable than when there are only spatially-varying slopes.

We introduce spatial dependence in the spatially varying slopes b_{1i}s via a conditional autoregressive (CAR) (Besag et al., 1991) prior. The CAR model, which provides a conditional and local specification of an intrinsic Gaussian Markov random field (GMRF), has been widely used to account for dependence in spatial areal data, not only in imaging but also in ecology, epidemiology, etc. (Rue and Held, 2005). Its widespread use is closely linked to its ability to provide immediate posterior full conditional distributions since it is specified as a set of full conditionals. This feature makes it computationally convenient for the Gibbs sampling algorithm and the more general Markov Chain Monte Carlo (MCMC) methods (Banerjee et al., 2014), especially for high dimensional spatial data. In a CAR model the spatial dependence in spatial areal data is characterized through a pre-specified neighborhood system.

In our application, we assume a first order neighborhood system. If two voxels, i and j, share a common face, we say that they are neighbors, denoted by $i \sim j$.

An $N \times N$ adjacency matrix $\boldsymbol{W}$ is created based on this neighborhood system, with (i,j)th entry $w_{ij} = 1$ if $i \sim j$ and $i \neq j$, and 0 otherwise (note that a voxel is not a neighbor of itself). Then, the CAR model on $\boldsymbol{b}_1 = (b_{11}, b_{12}, ..., b_{1N})^T$ is defined by the set of full conditionals:

$$\left[b_{1i} \mid \boldsymbol{b}_{1(-i)}, \sigma^2, \boldsymbol{W}\right] \sim \mathrm{N}\left(\sum\nolimits_j (w_{ij}/w_{i+})\, b_{1j},\ \sigma^2/w_{i+}\right), \tag{4.1}$$

where $w_{i+} = \sum_{j=1}^{N} w_{ij}$, the number of neighbors of voxel i. By Brook's Lemma (Brook, 1964), the full conditionals in Equation (4.1) uniquely define the joint prior on $\boldsymbol{b}_1$ up to a normalizing constant.

To complete the prior distribution specification, we assume flat, improper priors on β_0 and β_1 and place a conjugate prior IG(1, 1) on the variance parameters, τ^2 and σ^2, where IG denotes the inverse gamma distribution. Given the large number voxels, the prior information on the variance parameters is relatively weak.

Taken together, our proposed model is as follows:

$$\begin{aligned}
Y_i &= \mu_i + \epsilon_i \\
\mu_i &= \beta_0 + (\beta_1 + b_{1i}) X_i \\
\pi(\boldsymbol{\beta}) &= \pi \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} \propto 1 \\
\left[b_{1i} \mid \boldsymbol{b}_{1(-i)}, \sigma^2, \boldsymbol{W}\right] &\sim \mathrm{N}\left(\sum\nolimits_j \frac{w_{ij}}{w_{i+}}\, b_{1j},\ \sigma^2/w_{i+}\right), \quad i = 1, ..., N \\
\epsilon_i &\overset{iid}{\sim} \mathrm{N}(0, \tau^2) \\
\tau^2 &\sim \mathrm{IG}(1, 1) \\
\sigma^2 &\sim \mathrm{IG}(1, 1)
\end{aligned} \tag{4.2}$$

4.2.3 Loss Function

Voxels are classified into three classes: activated voxels, deactivated voxels and null voxels (neither activated nor deactivated). The three classes of voxels are defined as such to indicate whether the Blood-Oxygenation-Level-Dependent (BOLD) signal (Ogawa et al., 1992) of the associated voxel increases, decreases or does not change due to the stimulus relative to the baseline during an fMRI experiment. These classes are assigned based on whether the Z-statistic is above some upper threshold (activated), below some lower threshold (deactivated) or in between the lower and upper threshold (null). It is impossible to classify every voxel into the correct class. If an incorrect decision is made, a voxel is called a false negative if it is truly activated or deactivated but falsely classified as null. On the other hand, a voxel is called false positive if it is truly null but falsely classified as either activated or deactivated.

We use a Bayesian decision theoretical approach designed for pre-surgical fMRI data analysis, proposed by Liu et al. (2016), to classify voxels into these three classes. Unlike the standard classification procedure for fMRI data analysis which aims to control the type I error, i.e., preventing false positives, the proposed loss function allows the control of both false positives and false negatives.

The loss function proposed by Liu et al. (2016) is built upon the work of Müller et al. (2007) and is defined as

$$\begin{aligned} \mathrm{L}(\mathbf{m}, \boldsymbol{\delta}) \;&=\; \sum\nolimits_i \{-f(m_i)\delta_i - [1 - f(m_i)](1 - \delta_i) + k_1 f(m_i)(1 - \delta_i) \\ &\quad + k_2[1 - f(m_i)]\delta_i + t\delta_i\} \end{aligned}$$

where $\delta_i \in \{0, 1\}$ denotes the allocated class for voxel i, with 0 for null and 1 for non-null. The function f is chosen to be $f(m_i) = m_i/q_\alpha(\mathbf{m})$ based on Gross and Binder (2014) where q_α is the $(1 - \alpha)$ quantile function of its argument, $m_i = \left|\mathbb{E}(\mu_i \mid \boldsymbol{Y})/\sqrt{\mathrm{Var}(\mu_i \mid \boldsymbol{Y})}\right|$, $\mathbf{m} = (m_1, m_2, \ldots, m_N)$, and k_1, k_2 and t are positive

weights that control the number of false negatives, false positives, and total number of non-null voxels, respectively.

The optimal decision rule is given by

$$\delta_i = \mathrm{I}\left[f(\widehat{m}_i) \geq (1 + k_2 + t)/(2 + k_1 + k_2)\right],$$

where $\widehat{m}_i$ is the posterior estimate of m_i. By setting k_1 larger than k_2, a larger penalty is given to false negatives than false positives, resulting in relatively stricter control of false negatives—which is desirable for pre-surgical fMRI data analysis.

4.3 Simulation Study

In this section, we conduct simulation studies to evaluate the performance of the proposed model from three different aspects. First, we assess the ability of the proposed model to recover the parameters. Second, we examine its performance under different SNRs in both the high and low resolution images. Third, we compare the false negative rate (FNR) and false positive rate (FPR) of the proposed model with that of the model proposed by Liu et al. (2016), which is only applicable to first-level fMRI data at one given resolution. In what follows, we refer to our proposed model as the SVC model (spatially varying coefficients model) and the model proposed by Liu et al. (2016) as the CWAS model.

We simulate low resolution 4-dimensional fMRI data using the neuRosim package in R as described in Welvaert et al. (2011). We use the same brain mask as that of the patient for the real data application (see Section 4.4 for details). This brain mask image has a dimension of $64 \times 64 \times 48$ in low-resolution space and is used to indicate which voxels are in the brain and which ones are outside of the brain. Repetition time (TR) is 2 seconds and the total number of scans is 100. The stimulus function has a box-car block design with on and off periods each lasting 20 seconds

and is convolved with a gamma hemodynamic response function (HRF). Six activated regions are randomly selected at different locations and with different effect sizes and volumes (See Table 4.1 and row 1 of Figure 4.1 for details). The baseline signal is set to 250 and the signal fading rate is set to 0.01—values used by Welvaert et al. (2011). A mixture of Gaussian white noise (30%) and spatial noise (70%) is added to the signal. The amount of noise for the simulated low resolution fMRI data is controlled by the SNR= $\bar{S}/\sigma_N$, where $\bar{S}$ is the average magnitude of the signal, and σ_N is the standard deviation of the noise (Krüger and Glover, 2001). We choose three different values for the SNR, 1, 2, and 3. We process the 4-dimensional fMRI data into the 3-dimensional Z-statistic fMRI image using the FEAT function in the FSL software library without spatial smoothing. The interpolated low resolution image $\boldsymbol{X}$ is obtained by interpolating the unsmoothed Z-statistic image at low resolution $64 \times 64 \times 48$ to a higher resolution $120 \times 120 \times 62$ using the flirt function in FSL with nearest neighbor interpolation. In the simulated image $\boldsymbol{X}$, there are a total of 169047 voxels, of which 1943 are activated.

We simulate the high resolution image $\boldsymbol{Y}$ based on our model:

$$Y_i = \mu_i + \epsilon_i = \beta_0 + (\beta_1 + b_{1i})X_i + \epsilon_i$$

where $\epsilon_i \overset{iid}{\sim} \mathrm{N}(0, \tau^2)$, and, without loss of generality, we let $\beta_0 = 0$, $\beta_1 = 0$. We generate $\boldsymbol{b}_1$ from a Gaussian Random Field (GRF) with mean 0, standard deviation 0.4, and a power exponential correlation function. This class of correlation functions is specified as

$$\mathrm{Corr}(b_{1i},\, b_{1j}) = \exp\{-\theta|i-j|^p\}, \tag{4.3}$$

where $\theta > 0$ controls the degree of correlation between location i and j, and $p \in (0, 2]$ is a smoothing parameter (Abrahamsen, 1997). When $p = 2$ Equation 4.3 becomes the Gaussian, or squared-exponential, correlation function, while when $p = 1$ Equation

4.3 specifies the exponential correlation function. In our simulation experiments, we set $\theta = 400$ and $p = 1.99$ to generate a relative smoothing GRF for $\boldsymbol{b}_1$ but still with obvious peaks and dips (see row 1 of Figure 4.2 for an example). Finally, we select four different values for τ^2, i.e., $\tau^2 = 2, 4, 9, 25$, which governs the amount of noise added to the true μ_is. Representative slices of the simulated low resolution image, the interpolated low resolution image $\boldsymbol{X}$, and the high resolution image $\boldsymbol{Y}$ are shown in rows 2-4 of Figure 4.1, respectively.

With three different choices for the SNR of the low resolution image and with four different choices for the variance of the measurement error added to the smoothed Z-statistic image used to generate the high resolution image, we create 12 different scenarios that we use to examine the performance of the SVC model. Under each of the 12 scenarios, we simulate $M = 50$ data sets to which we fit the SVC model. On the other hand, we fit the CWAS model only to the simulated high resolution simulated images. For each model, we run the MCMC algorithm for 25,000 iterations with the first 10,000 iterations discarded as burn-in. Convergence is verified using the Gelman-Rubin convergence diagnostic statistic (Gelman and Rubin, 1992; Brooks and Gelman, 1998). After model fitting, we apply the Bayesian decision procedure described in Section 4.2.3 to determine activated regions.

Figure 4.2 compares the true values and the posterior mean estimates of the random slopes b_{1i}s and mean intensities μ_is under a specific scenario, SNR=2 and $\tau^2 = 4$. It is evident from this figure that the posterior mean estimates of $\boldsymbol{\mu}$ obtained by the SVC model are much closer to the truth than those obtained by the CWAS model. Additionally, as Figure 4.2 indicates, the spatial pattern of the true b_{1i}s is well captured by the posterior means of the $\mathbf{b}_1$ in the SVC model. Similar results are obtained under the other scenarios. We also note that the amount of smoothing differs between the SVC and the CWAS model (see Figure 4.4) .

We measure the ability of both models to recover the truth based on the average

mean square error ($\overline{\text{MSE}}$) and the average bias ($\overline{\text{Bias}}$) across all voxels:

$$\begin{aligned}\overline{\text{MSE}} &= (NM)^{-1}\sum_{i=1}^{N}\sum_{s=1}^{M}(\hat{\mu_i}^s - \mu_i)^2,\\ \overline{\text{Bias}} &= (NM)^{-1}\sum_{i=1}^{N}\sum_{s=1}^{M}(\hat{\mu_i}^s - \mu_i),\end{aligned}$$

where $\hat{\mu_i}^s$ is the posterior mean of μ_i for voxel i in the sth simulation. The results of $\overline{\text{MSE}}$ and $\overline{\text{Bias}}$ are summarized in Table 4.2 which shows that the SVC model is more accurate than the CWAS model under all 12 scenarios.

Lastly, we assess the performance of both models based on FNR and FPR. Figure 4.3 shows the FNR and the FPR versus threshold under 12 scenarios for both the SVC and the CWAS models. For each scenario, we take the average of the FNR or the FPR over 50 simulations. We let the threshold vary from 0 to 1 instead of fixing it at a specific point to investigate the trade-off between the FNR and FPR as a function of the threshold. Figure 4.3 shows that both the FNR and FPR do not change as much under the SVC with increasing amounts of noise, i.e., with increasing τ^2, demonstrating that the SVC model is more robust to changes in the SNR. Furthermore, the SVC model has consistently lower FNR and FPR than the CWAS model for all threshold settings. Through these simulation studies, we can conclude that the SVC model outperforms the CWAS model by exploiting the information contained in both the low and high resolution fMRI data, achieving a balance between high spatial precision and high SNR.

4.4 Pre-surgical fMRI Data Analysis

In this section, we apply our method to a pre-surgical fMRI data acquired at two different spatial resolutions from a 62 year old, right-handed woman with a glioblastoma of the left middle-to-inferior posterior temporal lobe transferred for pre-surgical

language mapping. We compare the results obtained from our proposed model with those obtained from the CWAS model (Liu et al., 2016). As the patient had a tumor located near the so-called ventral stream of speech and language processing (Hickok and Poeppel, 2007) and showed some difficulties in reading and comprehension, the fMRI experiment was designed to challenge functions of the ventral stream. Specifically, a 30 seconds ON/OFF boxcar block design was implemented and the patient was asked to silently alternate between reading non-finite embedded clause sentences (ON) versus consonant strings (OFF) that were ordered like sentences. The patient was not aphasic and demonstrated no agraphia or acalculia but exhibited video-documented color anomia consistent with the lesion location. She performed the fMRI experiment at both low and high spatial resolutions. The acquired fMRI data set at the low spatial resolution has a total number of $64 \times 64 \times 48$ voxels in space with voxel dimensions of $3.00 \times 3.00 \times 3.45$ mm^3, while the higher spatial resolution has a total of $120 \times 120 \times 62$ voxels in space, with voxel dimensions of $1.80 \times 1.80 \times 2.30$ mm^3. The temporal resolution was identical for both the low and high spatial resolution data sets and contained 160 volumes in time. The low-resolution fMRI data was registered to high-resolution space using the registration method described in Section 4.2.1. The FSL software was used to perform pre-processing (no spatial smoothing) and mass univariate analysis of the data, generating Z-statistic images.

Data were acquired at 3 Tesla (Magnetom TimTrio scanner, Siemens Healthcare, Erlangen, Germany), using a 32 channel head coil and identical repetition (TR = 3000ms) and echo times (TE = 30ms) of gradient echo, T2*-weighted BOLD-sensitive echo-planar imaging (EPI). We took advantage of integrated parallel acquisition technique (iPAT) by using Generalized Autocalibrating Partially Parallel Acquisition (GRAPPA) at an acceleration factor of 2, resulting in an effective echo spacing of 0.345 ms. We applied the following pre-statistics processing: motion correction using MCFLIRT (Jenkinson et al., 2002), non-brain removal using BET (Smith, 2002),

grand-mean intensity normalisation of the entire 4D dataset by a single multiplicative factor and highpass temporal filtering (Gaussian-weighted least-squares straight line fitting, with sigma=45.0 secs, resulting in a high pass filter cutoff of approximately 90 secs). We performed distortion correction using dual (TE = 4.92 and 7.38 ms, TR = 456 ms) gradient echo fieldmaps coplanar to the fMRI data and PRELUDE/FUGUE (both part of FSL). Finally, Slices were acquired in sagittal planes and at an inter-slice gap of 15% the slice thickness (3 mm for the low resolution data, 2 mm for the high resolution data). For the high resolution data, we used simultaneous multi-slice imaging (SMS) with an acceleration factor of 2.

We fit the SVC model (Equation 4.2) using the co-registered low- and high-resolution Z-statistic images, while we fit the CWAS model to the Z-statistic images individually, e.g. to the high resolution Z-statistic image and to the low resolution Z-statistic image that didn't undergo interpolation, separately. To determine the activated and deactivated regions of the brain for each model, we apply the loss function described in Section 4.2.3 proposed by Liu et al. (2016). In the loss function, we use the quantile level $\alpha = 0$, giving $f(m_i) = m_i / \max(\mathbf{m})$, and a threshold of 0.2, that results in false negatives being penalized 11 times more than false positives. This implies that any voxel with an estimated posterior mean $//\mathbb{E}(\hat{\mu}_i \mid \boldsymbol{Y}) > 0$ greater than zero and $f(m_i) = m_i / q_0(\mathbf{m}) \geq 0.2$ is classified as an activated voxel while any voxel with estimated posterior mean less than zero $//\mathbb{E}(\mu_i \mid \boldsymbol{Y}) < 0$ and $f(m_i) = m_i / q_0(\mathbf{m}) \geq 0.2$ is classified as deactivated voxel.

We run our model for 50,000 iterations, with the first 10,000 iterations discarded as burn-in. The Gibbs sampling algorithm takes approximately 30 mins CPU time on an iMAC with a 3.2 GHz Intel Core i5 processor and a 24 GB memory. To assess convergence of the MCMC chains for all parameters, we use the univariate version of the Gelman-Rubin R statistic (Gelman and Rubin, 1992; Brooks and Gelman, 1998) instead of the multivariate version due to the very large number of parameters

(over 150k parameters). We obtain the Gelman-Rubin R statistic from five MCMC simulations, each run with different initial values. The range of the Gelman-Rubin statistic for the b_{1i}s and the Gelman-Rubin statistics for τ^2, σ^2, β_0 and β_1 are provided in Table 4.3 along with the MCMC standard errors, computed using the "mcmcse" package in R (Flegal and Hughes, 2012) with the default batch means method (Flegal et al., 2008). Since all of the Gelman-Rubin R statistics are below the cutoff 1.2, we confirm convergence of the MCMC chains.

We examine the smoothing properties of the SVC model and the CWAS model via scatter plots of the estimated posterior means versus the unsmoothed image. Figure 4.4(a) displays such a scatter plot of the estimated posterior means generated by the SVC model versus the unsmoothed high resolution image (Y_i). Figures 4.4 (b) and (c) display the scatter plots of the estimated posterior means generated by the CWAS model versus the unsmoothed high resolution image (Y_i) and the unsmoothed low resolution image (X_i), respectively. The differentiable influences of the low resolution image on the smoothing degree, as indicated by the color coding of whether X_i (low-resolution image voxel value) is within (-2, 2) (Figure 4.4(a)), reveals that the SVC model is able to adaptively smooth the high resolution image by incorporating the low resolution image information, whereas the CWAS model utilizes only information from a single spatial resolution and is thus less informative than the SVC model.

Maximum intensity projection views of the results from the CWAS and SVC models on the loss function scale are shown in Figure 4.5. This figure indicates that the SVC model shows the best functional contrast-to-noise ratio in the statistic image, probably as a result of the combination of information from the two data sets that leads to a preservation of the significant regions identified by the low resolution image while preventing the random noise from the high resolution data to blur the signal. Figure 4.6 presents a comparisons of the activated and deactivated regions as detected by the SVC and CWAS model on the truncated loss function scale, with the

regions projected onto the cortical surface. We use FreeSurfer (http://surfer.nmr.-mgh.harvard.edu/; Fischl (2012)) to construct cortical surface views of the activated and deactivated regions. Since brain activities may lead to a decrease in fMRI signal in patients with tumor in the brain (Fujiwara et al., 2004; Hsu et al., 2004; Ulmer et al., 2004; Bartsch et al., 2006), in presenting results of our analysis, we consider and discuss activated (red-to-yellow) and deactivated (blue-to-lightblue) regions together. As we note from Figure 4.6, the activated and deactivated regions detected by the CWAS model based on either the high resolution or low resolution image data are quite different from those detected by the SVC model which, in comparison, are much more localized. To interpret these results, one important aspect that cannot be ignored is that the fMRI data discussed in this analysis are real biological data and results might be affected by biological and run-to-run variability. In particular, as the patient first performed the experiment at low resolution, and then again at high resolution, there may be a learned confounding effect at play affecting the images. However, multiple runs of the same paradigm are generally considered beneficial in improving clinical fMRI results, and here two runs were recorded at different spatial resolutions. Furthermore, real data acquired at different spatial resolutions exhibit different T2*-decay related blurring – despite adequate correction of geometric susceptibility-induced distortions. This may also confound the results to some extent.

4.5 Discussion

In this paper, we present a modeling approach that combines information from pre-surgical fMRI data acquired at both high and low spatial image resolutions. The proposed mixed-effect model with a spatially varying slope is able to retain the spatial precision from the high resolution image and improve accuracy of the activation/deactivation map by leveraging data from the low resolution image. We use

a comparative analysis to evaluate the performance of the proposed model and an existing single-resolution CWAS model from Liu et al. (2016). The performance of the CWAS model is found to be sensitive to the SNR of the high resolution image. Conversely, the SVC model is robust to the SNR and consistently outperforms CWAS on activated/deactivated classification accuracy.

With respect to the clinical application of high resolution fMRI, we need to point out that it still remains controversial whether the spatial resolution of fMRI data should be pushed higher. Increasing the spatial resolution of fMRI may help to improve the spatial accuracy of the functional mapping. Pushing spatial resolution much higher may eventually allow imaging of details of the brain, such as cortical columns which are basic functional units that consist of groups of neurons in the cortex of the brain (Kim et al., 2000). However, increasing the spatial resolution comes at the cost of loss in the SNR and, eventually, a decreased sensitivity of fMRI to detect functionally activated regions (Yoo et al., 2001). The data recorded at low resolution have more signal while high resolution data are much more noisy and often require to be smoothed approximately by the same kernel sizes for fMRI data at standard resolution just to get rid of the extra noise (Harms et al., 2013). Additionally, high resolution of fMRI may not entail high specificity to the local neural activity. As the spatial resolution increases, the fMRI signal specificity suffers increasingly from functionally less-specific non-capillary signals (Boxerman et al., 1995; Siero et al., 2013).

One limitation of our model is its focus on fMRI data sets at two different resolutions. It would be interesting to generalize the model to incorporate multiple data sets (images) at varying resolutions. Studies can also be conducted to evaluate the improvements in classification accuracy with respect to the data size and the resolutions. Here we provide two ways on how the proposed model can be extended. First, the model may be extended to a multivariate version to incorporate multiple images

at each resolution. A simple case would be two low-resolution images and two high-resolution images. A more complicated case would be multiple low-resolution images and one high-resolution image. Second, the model may be extended to incorporate images with more than two resolutions. Since there is a trade-off between spatial resolution and SNR, it would be beneficial to combining all data sources for the best performance of estimation and classification. But it still remains a question as how to interpolate the data with more than two resolutions and how to deal with the possible co-linearity issue between two images (covariates).

The SVC model that we have introduced here provides an entirely new tool to evaluate the benefits of fMRI data recorded at two different (high and low) spatial image resolutions. Considering recent advancements in fMRI scanning, such as simultaneous multi-slice (SMS) and zoomed selective field-of-view imaging without aliasing artifacts (ZOOMit) (Riffel and Michaely, 2013) that allow to increase the spatial and/or temporal image resolution, our model may prove beneficial to optimize the gains that can be achieved by these technologies. By combining fMRI data from different resources, our method is particularly promising for pre-surgical fMRI applications.

Center Voxel	[37,44,11]	[26,34,12]	[52,32,21]	[40,45,26]	[28,38,31]	[41,38,40]
Effect Size	8	2	4	1	6	10
Radius	4	6	7	4	2	1

Table 4.1: Locations, effect sizes and radii of the six truly activated regions simulated in the low resolution image.

		SNR=1		SNR=2		SNR=3	
		SVC	CWAS	SVC	CWAS	SVC	CWAS
$\tau^2 = 2$	$\overline{MSE}(SD)$	0.055 (0.034)	0.779 (0.753)	0.057 (0.043)	0.780 (0.752)	0.058 (0.061)	0.781 (0.752)
	$\overline{Bias}(SD)$	-0.001 (0.036)	-0.001 (0.125)	-0.001 (0.037)	-0.001 (0.127)	-0.001 (0.036)	-0.001 (0.129)
$\tau^2 = 4$	$\overline{MSE}(SD)$	0.071 (0.049)	1.534 (1.517)	0.073 (0.059)	1.535 (1.516)	0.075 (0.084)	1.536 (1.515)
	$\overline{Bias}(SD)$	-0.001 (0.041)	-0.001 (0.175)	-0.001 (0.043)	-0.001 (0.176)	-0.001 (0.043)	-0.001 (0.178)
$\tau^2 = 9$	$\overline{MSE}(SD)$	0.093 (0.072)	3.418 (3.445)	0.096 (0.086)	3.419 (3.443)	0.099 (0.118)	3.419 (3.439)
	$\overline{Bias}(SD)$	0.001 (0.049)	0.001 (0.262)	0.001 (0.053)	0.001 (0.263)	0.001 (0.053)	0.001 (0.264)
$\tau^2 = 25$	$\overline{MSE}(SD)$	0.127 (0.121)	9.454 (9.656)	0.132 (0.136)	9.454 (9.651)	0.137 (0.174)	9.455 (9.647)
	$\overline{Bias}(SD)$	-0.002 (0.064)	-0.002 (0.435)	-0.002 (0.071)	-0.002 (0.435)	-0.002 (0.073)	-0.002 (0.436)

Table 4.2: Simulation results. Average mean square error ($\overline{\text{MSE}}$) and the average bias ($\overline{\text{Bias}}$) across all voxels from both the SVC model and the CWAS model.

	β_0	β_1	τ^2	σ^2	$b'_{1i}s$
Posterior Estimate	0.008	0.121	2.538	0.452	(-1.28, 1.38)
MCMC SE	3.57e-05	3.13e-05	2.31e-04	5.68e-04	(5.10e-04, 1.54e-02)
Gelman-Rubin Statistic	1.00004	1.00009	1.00071	1.00193	(0.99999, 1.00070)

Table 4.3: Posterior mean, MCMC standard error (SE) and Gelman-Rubin Statistic for all parameters. For the $b'_{1i}s$, we provide the range.

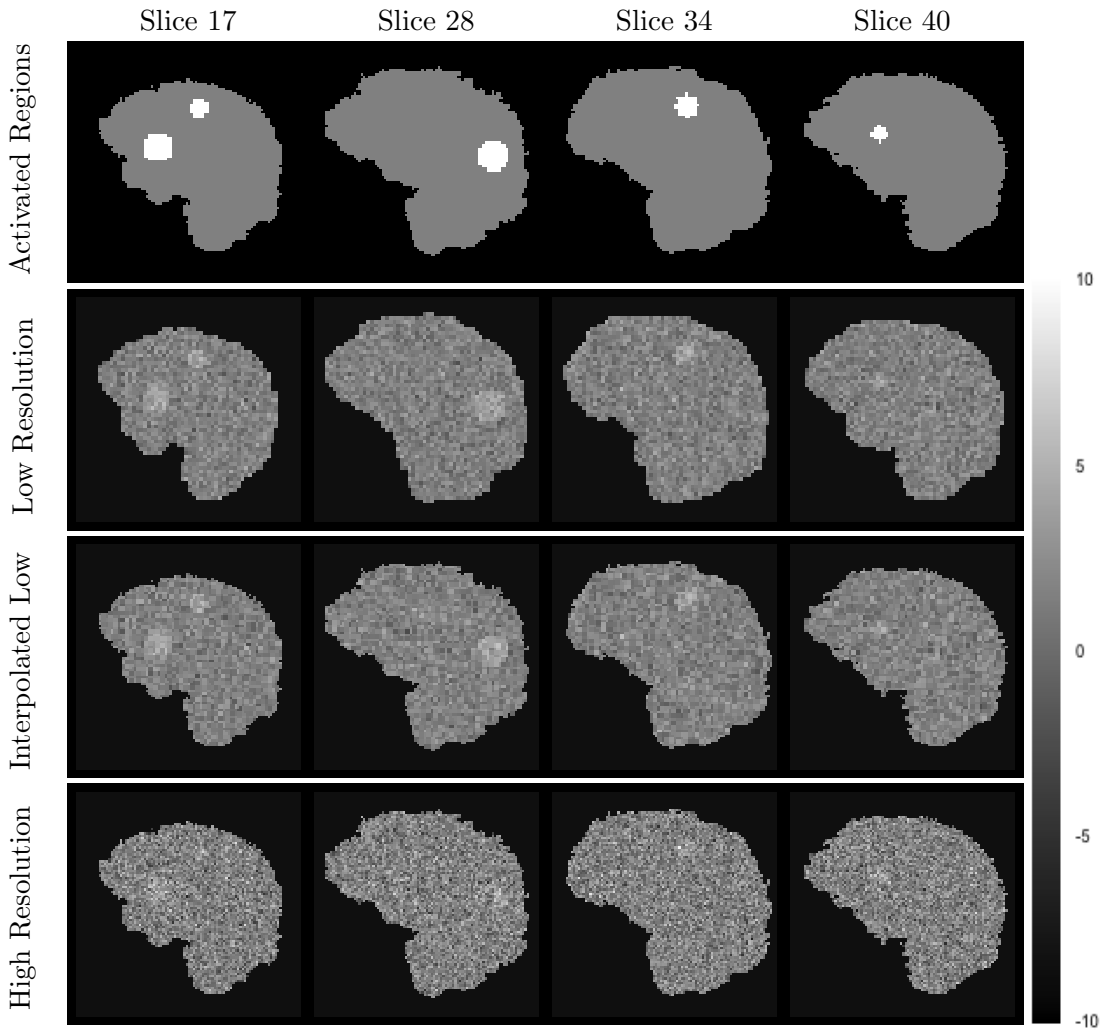

Figure 4.1: Representative sagittal slices of the simulated fMRI data when SNR = 2 and $\tau^2 = 4$. Row 1: Truly activated regions (one of the six simulated truly activated regions does not appear in these four slices). Row 2: Simulated low resolution Z-statistic image without smoothing. Row 3: Interpolated low resolution image with nearest neighbor interpolation method. Row 4: Simulated high resolution image (Note that the loss of SNR associated with high spatial resolution results in less clearly defined activations).

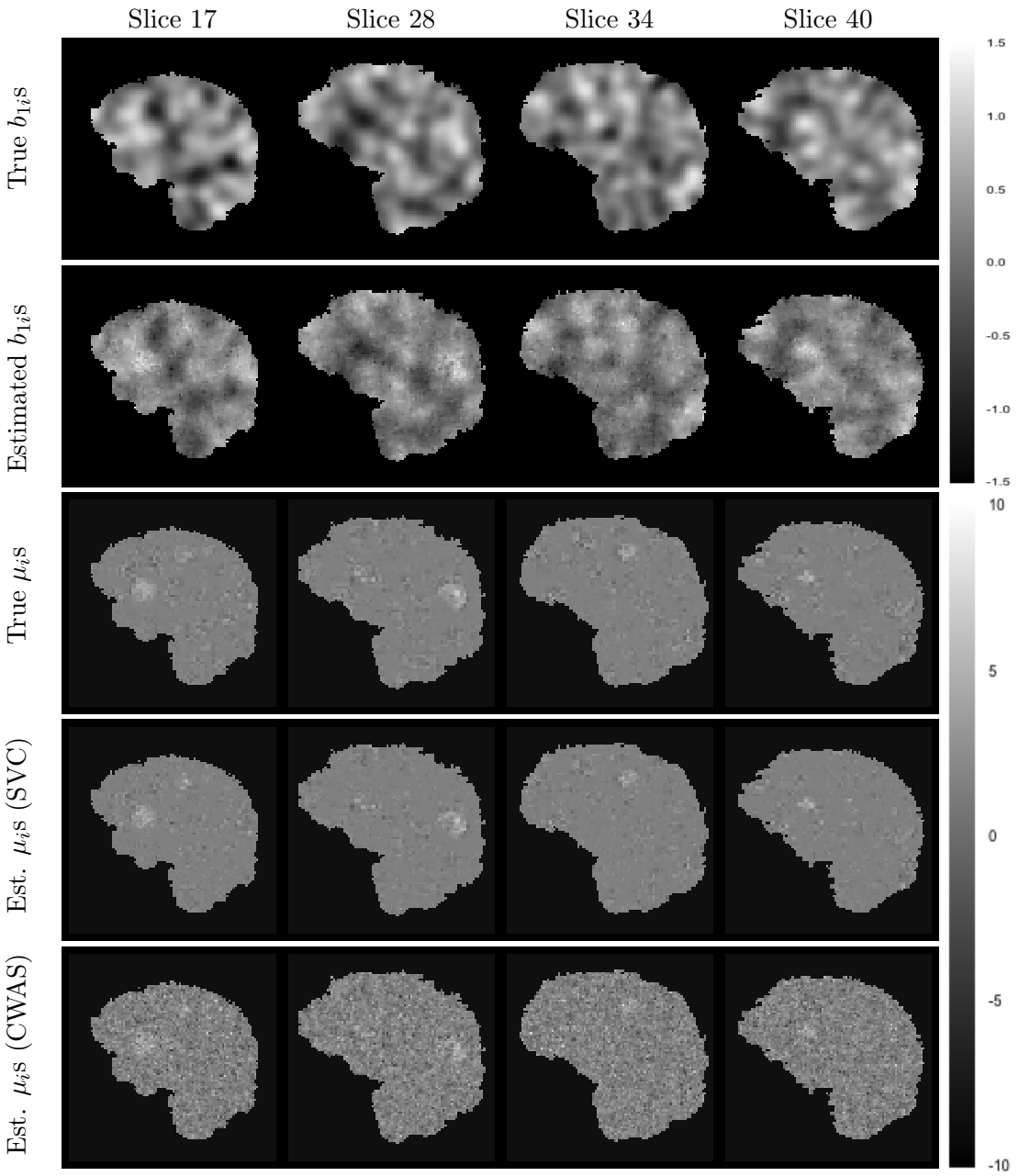

Figure 4.2: Comparison of the true values and the posterior means of the random slopes b_{1i}s and mean intensities μ_is when SNR = 2 and $\tau^2 = 4$. Row 1-2 are true b_{1i}s and the marginal posterior means of b_{1i}s from the SVC model,respectively. Row 3-5 are true μ_is, marginal posterior means of the μ_is from the SVC model and marginal posterior means of the μ_is from the CWAS model, respectively. Note that the SVC model retains a higher contrast-to-noise ratio compared to the CWAS model.

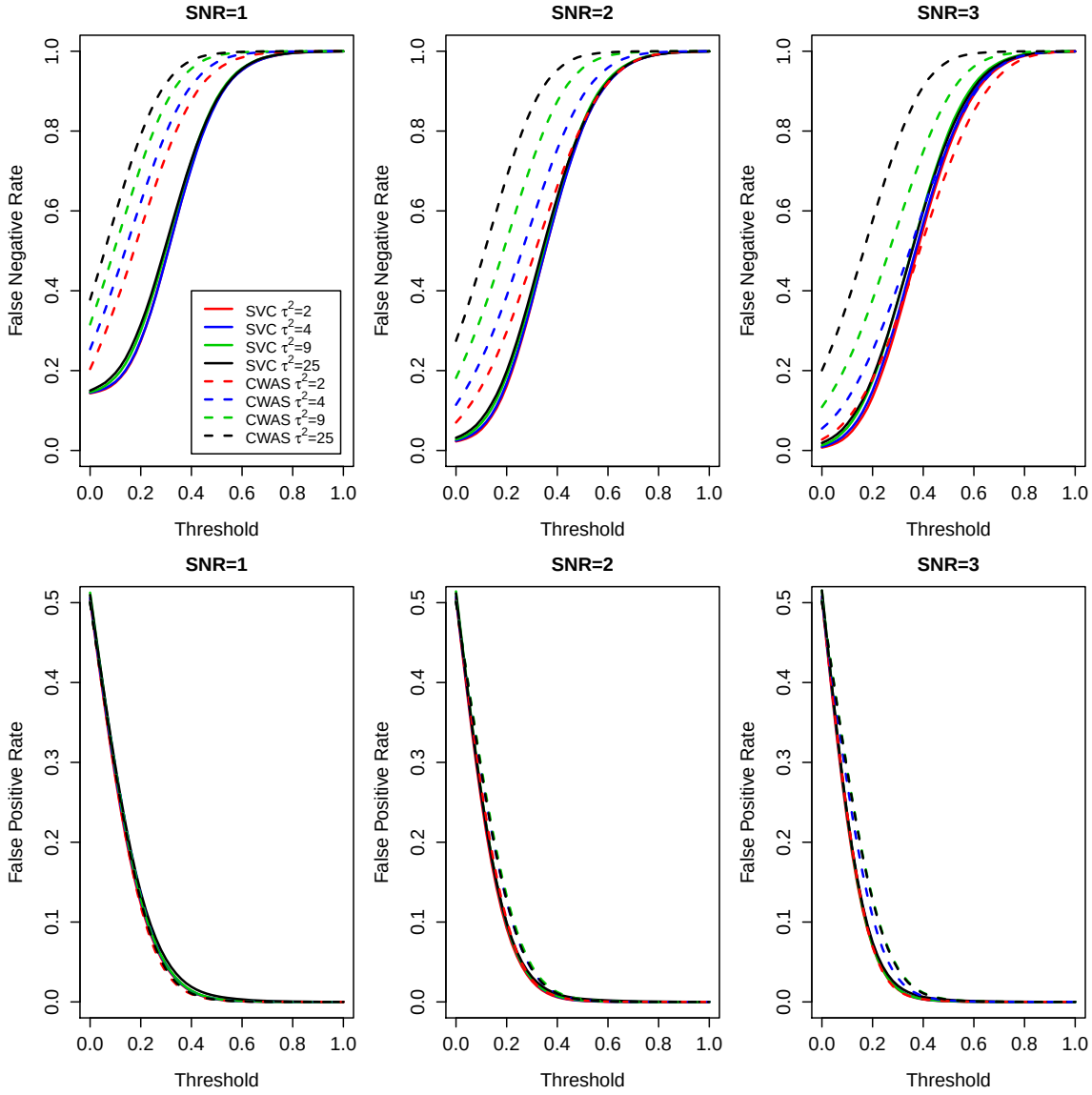

Figure 4.3: False negative rate (FNR) and false positive rate (FPR) vs threshold under 12 scenarios for both the SVC model and the CWAS model in the simulation studies. For each scenario, we take the average of the FNR or the FPR over 50 simulations. Threshold varies from 0 to 1. The legend in the first panel is for all six panels.

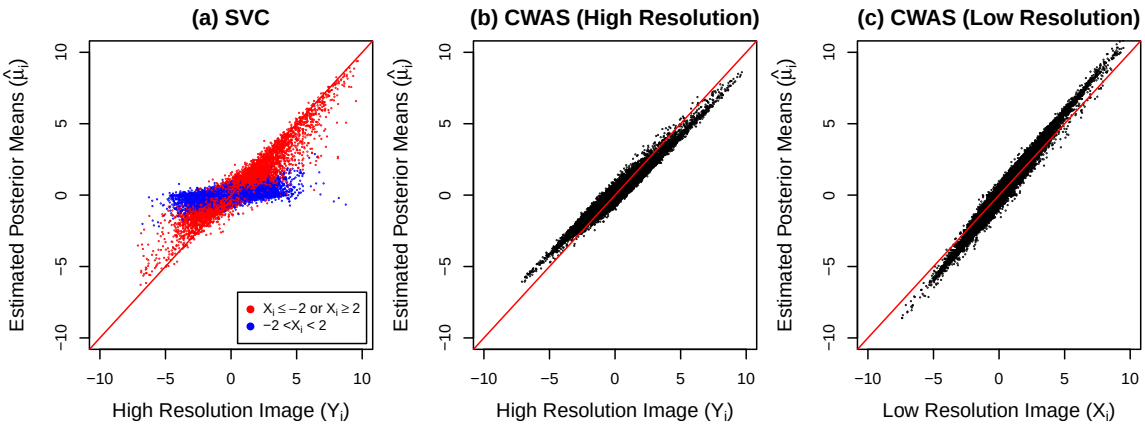

Figure 4.4: Scatter plots of the estimated posterior means versus the unsmoothed image. (a) The estimated posterior means ($\hat{\mu}_i$) generated by the SVC model versus the unsmoothed high resolution image (Y_i). (b) The estimated posterior means ($\hat{\mu}_i$) generated by the CWAS model versus the unsmoothed high resolution image (Y_i). (c) The estimated posterior means ($\hat{\mu}_i$) generated by the CWAS model versus the unsmoothed low resolution image (X_i).

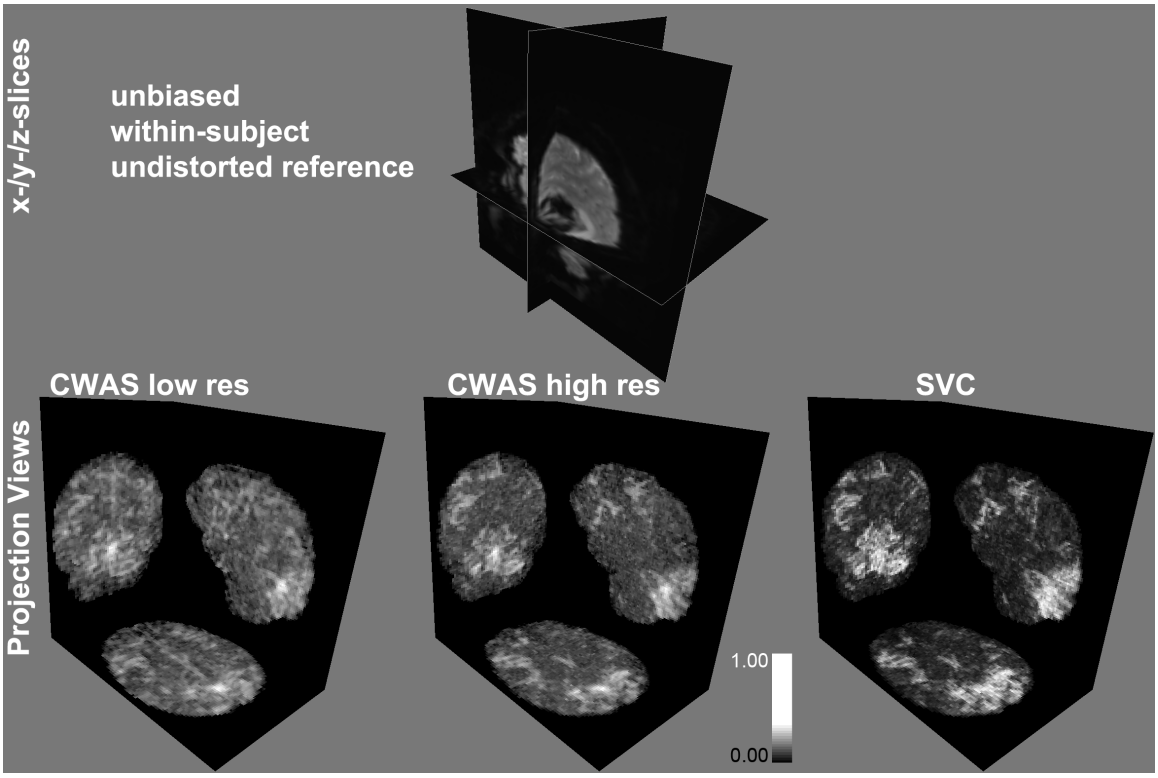

Figure 4.5: Maximum intensity projection views of the results from the CWAS and SVC models on loss function scale. The projection views use a maximum intensity projection scaled between 0 and 1. The best contrast-to-noise ratio is achieved by the SVC model.

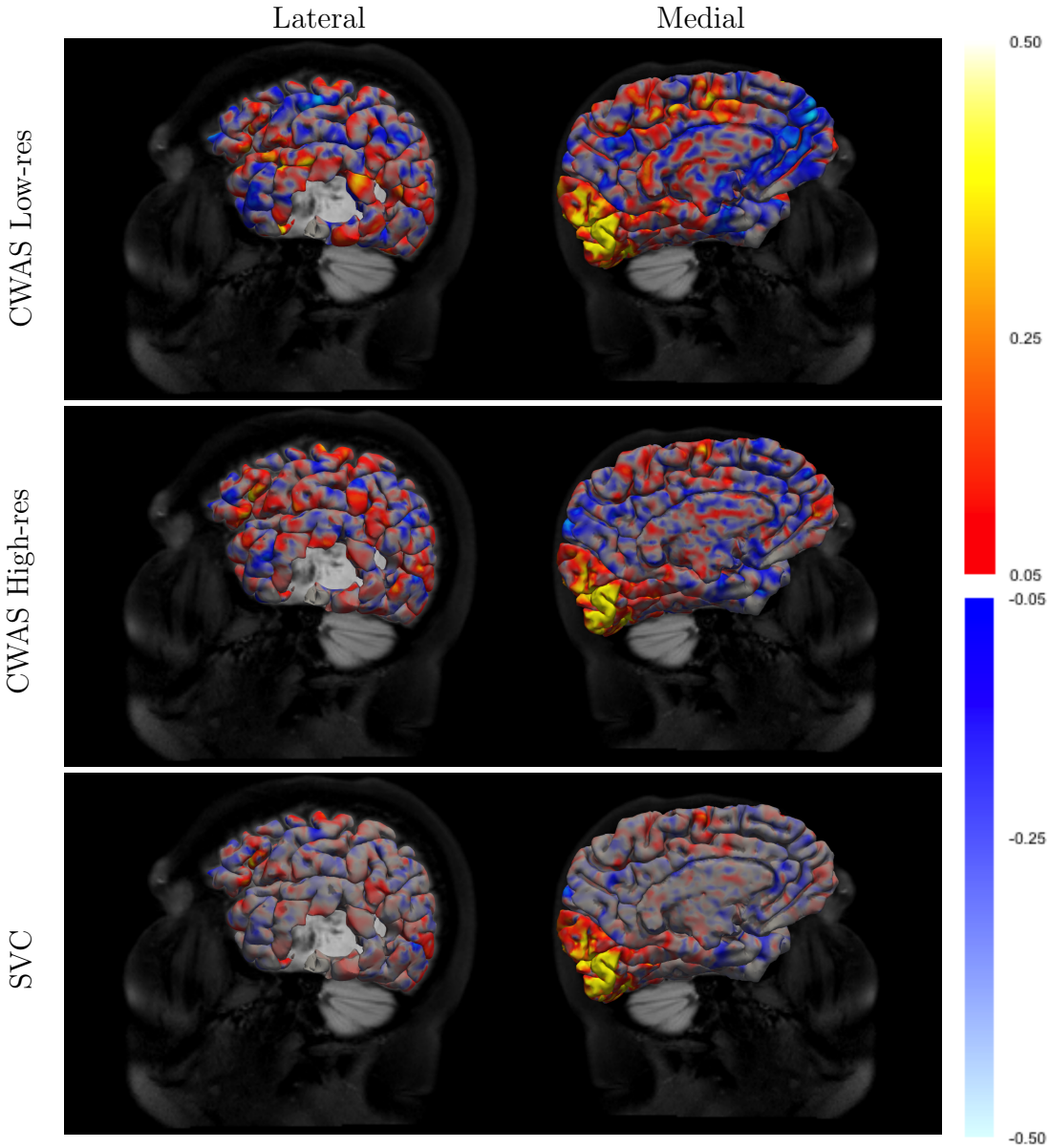

Figure 4.6: Surface projections of the results from the CWAS and SVC models on truncated loss function scale: loss function values from 0.05 to 0.5 indicate increasingly strong activation and loss function values form -0.05 to -0.5 indicated increasingly strong deactivation. To better differentiate the voxel intensities, values above 0.5 are mapped to 0.5 and values below -0.5 are mapped to -0.5.

4.6 Appendix

We implement the Gibbs sampling algorithm to estimate all the parameters of the proposed model in the Bayesian framework since the full conditionals for these parameters are all available in closed forms. The joint posterior distribution is:

$$\pi(\boldsymbol{\beta}, \boldsymbol{b}_1, \sigma^2, \tau^2 \mid \boldsymbol{Y}, \boldsymbol{X}) \propto L(\boldsymbol{Y} \mid \boldsymbol{X}, \boldsymbol{\beta}, \boldsymbol{b}_1, \tau^2)\pi(\boldsymbol{\beta})\pi(\boldsymbol{b}_1 \mid \sigma^2)\pi(\sigma^2)\pi(\tau^2)$$

The full conditionals for $\boldsymbol{\beta}$, b_{1i}, σ^2, τ^2 are multivariate normal distribution, normal distribution, inverse gamma distribution and inverse gamma distribution, respectively. The four steps of the Gibbs sampling algorithm is:

- Sampling $\boldsymbol{\beta}$ via

$$\pi(\boldsymbol{\beta}|\cdot) \sim \text{MVN}\left((\boldsymbol{Z}^T\boldsymbol{Z})^{-1}\boldsymbol{Z}^T(\boldsymbol{Y} - \boldsymbol{b}_1 \circ \boldsymbol{Z}),\, \tau^2(\boldsymbol{Z}^T\boldsymbol{Z})^{-1}\right)$$

  where $\boldsymbol{Z} = (\mathbf{1}^T, \boldsymbol{X}^T)^T$, $\mathbf{1}$ denotes a vector of 1 with length N and $\circ$ denotes the element wise matrix multiplication.

- Sampling b_{1i} via

$$\pi(b_{1i}|\cdot) \sim \text{N}\left(V\left(b^\star_{1i} w_{i+}/\sigma^2 + (y_i - \beta_0 - \beta_1 x_i)x_i/\tau^2\right),\, V\right).$$

  where $b^\star_{1i} = \sum_j \frac{w_{ij}}{w_{i+}} b_{1j}$ and $V = (w_{i+}/\sigma^2 + x_i^2/\tau^2)^{-1}$.

- Sampling σ^2 via

$$\pi(\sigma^2|\cdot) \sim \text{IG}\left(0.5(N-1)+1,\, 0.5\sum\nolimits_{i\sim j}(b_{1i} - b_{1j})^2 + 1\right).$$

- Sampling τ^2 via

$$\pi(\tau^2|\cdot) \sim \text{IG}\left(0.5N + 1,\, 0.5\sum\nolimits_i (y_i - (\beta_0 + (\beta_1 + b_{1i})x_i))^2 + 1\right).$$

CHAPTER V

Discussion

In this dissertation, we developed novel statistical modeling and inference methods for fMRI data in the Bayesian framework. In particular, we focused on applying the proposed methods to the analysis of pre-surgical fMRI data in the context of surgery for patients with brain tumors.

In Chapter II, we proposed a novel spatially adaptive conditionally autoregressive model, i. e., the CWAS model, that adaptively and locally smooths the fMRI data, and a Bayesian theoretical decision approach that accounts for the different consequences false positives and false negatives incurred in the identification of activated and deactivated brain regions. Our model provides more flexibility in characterizing the spatial variation of the fMRI data and introduces a parameter which offers an intuitive interpretation in controlling the degree of smoothing. Further, the proposed model that utilizes pseudo-likelihood approximation for Bayesian posterior simulation is shown to be effective in dealing with high-dimensional fMRI data.

In Chapter III, we developed a series of models which extend the idea of spatially adaptive smoothing to multiple fMRI datasets. In particular, we propose three spatially adaptive multivariate conditional autoregressive (SAMCAR) models that can be considered as extensions of the multivariate conditional autoregressive (MCAR) model (Carlin and Banerjee, 2003; Gelfand and Vounatsou, 2003), the CWAS model

(Liu et al., 2016) and the RH model (Reich and Hodges, 2008), respectively, and one mixed-effects model assuming that all observed fMRI images originate from one common image. All proposed models are designed to characterize spatial correlations not only within images but also across multiple images to improve inference. We evaluated the performance of all models using both simulated independent images and pair of images simulated with different pre-defined correlations. The models are applied to two sets of fMRI brain images, acquired either from the same patient, same paradigm or same patient, different paradigms. The results indicate that the multivariate extension of the RH model is able to adaptively recover the varying correlations between the images.

In Chapter IV, motivated by fMRI datasets acquired at two different resolutions from the same patient, we developed a Bayesian hierarchical model with spatially varying coefficients to incorporate information from both high- and low-resolution data. The commonly used standard spatial resolution is relatively low but less noisy while the high resolution is associated with lower SNR due to the decrease in the strength of the fMRI signal as the voxel size is smaller. The proposed model leverages the high SNR from the low resolution data and retains the high spatial precision from the high resolution data. Specifically, we first interpolate the low-resolution image into the high-resolution space using nearest-neighbor interpolation and then we regress the high-resolution image on the low-resolution image via a mixed-effects model with spatially varying slopes. The results show that the proposed model is robust to the SNR of the high-resolution image and outperforms the CWAS model on activation/deactivation classification accuracy.

For future work, there are several possible directions. First, we can extend the proposed models to incorporate different types of priors to specify the local spatial dependencies in the fMRI data. In the proposed models, we adopted the conditional autoregressive (CAR) model as priors on the mean voxel intensities or the logarithm

of the variances. One alternative would be to use Laplace pairwise difference model (Besag, 1993) which is defined by the set of full conditionals: $\mu_{(-i)}$ is:

$$\pi(\mu_i \mid \mu_{(-i)}, \tau) \propto \frac{1}{\tau}\exp\left\{-\frac{1}{\tau}\sum_{j\in\partial i} \mid \mu_i - \mu_j \mid\right\}, \tag{5.1}$$

where ∂i denotes the neighbors of voxel i. The difference between the Laplace pairwise difference model and the CAR model, also called Gaussian pairwise difference model, is that the former model's mode is at the median whereas the later one is at the mean. The Laplace pairwise difference model is therefore more appropriate if the activated/deactivated regions have sharp boundaries.

Second, we can extend the proposed spatially adaptive models to incorporate spatial information about the tumor to improve the performance of the models in terms of identifying activated/deactivated regions, especially the peri- and intra-tumoral activated regions. Potentially useful spatial information includes, but is not limited to, the distance between the voxel and the edge or center of the tumor and whether the voxel is in grey matter or white matter. This information could be used as covariate information.

Third, it would be interesting to investigate loss functions with different choices of $f(m_i)$ to optimize the control of false positives and false negatives. For example, an alternative choice of $f(m_i)$ is the complementary cumulative distribution function for the normal distribution with mean $\mathbb{E}(\mu_i \mid \boldsymbol{Y})$ and variance $\mathrm{Var}(\mu_i \mid \boldsymbol{Y})$ evaluated at a pre-determined constant c_0, i.e., $f(m_i) = 1 - F_X(x = c_0)$, where $X \sim \mathrm{N}(\mathbb{E}(\mu_i \mid \boldsymbol{Y}), \mathrm{Var}(\mu_i \mid \boldsymbol{Y}))$. This specification is particular useful to guard against outliers caused by extremely small posterior variances of μ_i, i.e., $\mathrm{Var}(\mu_i \mid \boldsymbol{Y})$. Another interesting modification of the loss function is to convert the false positive and false negative counts to false positive and false negative rates, respectively.

BIBLIOGRAPHY

BIBLIOGRAPHY

Abrahamsen, P. (1997), *A review of Gaussian random fields and correlation functions*, Norsk Regnesentral/Norwegian Computing Center.

Banerjee, S., B. P. Carlin, and A. E. Gelfand (2014), *Hierarchical modeling and analysis for spatial data*, Crc Press.

Bartsch, A. J., G. Homola, A. Biller, L. Solymosi, and M. Bendszus (2006), Diagnostic functional MRI: Illustrated clinical applications and decision-making, *Journal of Magnetic Resonance Imaging*, *23*(6), 921–932.

Bartsch, A. J., A. Biller, and G. Homola (2014), Presurgical tractography applications, in *Diffusion MRI: from quantitative measurement to in-vivo neuroanatomy*, edited by H. Johansen-Berg and T. E. Behrens, Second ed., pp. 531–568, Elsevier Academic Press.

Beckmann, C. F., and S. M. Smith (2004), Probabilistic independent component analysis for functional magnetic resonance imaging, *Medical Imaging, IEEE Transactions on*, *23*(2), 137–152.

Beckmann, C. F., M. Jenkinson, and S. M. Smith (2003), General multilevel linear modeling for group analysis in FMRI, *Neuroimage*, *20*(2), 1052–1063.

Berrocal, V. J., A. E. Gelfand, and D. M. Holland (2010), A spatio-temporal downscaler for output from numerical models, *Journal of Agricultural, Biological, and Environmental Statistics*, *15*(2), 176–197.

Besag, J. (1974), Spatial interaction and the statistical analysis of lattice systems, *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 192–236.

Besag, J. (1975), Statistical analysis of non-lattice data, *The Statistician*, pp. 179–195.

Besag, J. (1977), Efficiency of pseudolikelihood estimation for simple Gaussian fields, *Biometrika*, pp. 616–618.

Besag, J. (1993), Towards Bayesian image analysis, *Journal of Applied Statistics*, *20*(5-6), 107–119.

Besag, J., J. York, and A. Mollié (1991), Bayesian image restoration, with two applications in spatial statistics, *Annals of the Institute of Statistical Mathematics*, *43*(1), 1–20.

Bookheimer, S. (2007), Pre-surgical language mapping with functional magnetic resonance imaging, *Neuropsychology Review*, *17*(2), 145–155.

Boxerman, J. L., L. M. Hamberg, B. R. Rosen, and R. M. Weisskoff (1995), MR contrast due to intravascular magnetic susceptibility perturbations, *Magnetic Resonance in Medicine*, *34*(4), 555–566.

Brewer, M. J., and A. J. Nolan (2007), Variable smoothing in Bayesian intrinsic autoregressions, *Environmetrics*, *18*(8), 841–857.

Brook, D. (1964), On the distinction between the conditional probability and the joint probability approaches in the specification of nearest-neighbour systems, *Biometrika*, *51*(3/4), 481–483.

Brooks, S. P., and A. Gelman (1998), General methods for monitoring convergence of iterative simulations, *Journal of Computational and Graphical Statistics*, *7*(4), 434–455.

Brown, R. W., Y.-C. N. Cheng, E. M. Haacke, M. R. Thompson, and R. Venkatesan (2014), *Magnetic resonance imaging: physical principles and sequence design*, John Wiley & Sons.

Bullmore, E., M. Brammer, S. C. Williams, S. Rabe-Hesketh, N. Janot, A. David, J. Mellers, R. Howard, and P. Sham (1996), Statistical methods of estimation and inference for functional MR image analysis, *Magnetic Resonance in Medicine*, *35*(2), 261–277.

Carlin, B. P., and S. Banerjee (2003), Hierarchical multivariate CAR models for spatio-temporally correlated survival data, *Bayesian Statistics*, *7*, 45–63.

Durnez, J., B. Moerkerke, A. Bartsch, and T. E. Nichols (2013), Alternative-based thresholding with application to presurgical fMRI, *Cognitive, Affective, & Behavioral Neuroscience*, *13*(4), 703–713.

Feinberg, D. A., et al. (2010), Multiplexed echo planar imaging for sub-second whole brain FMRI and fast diffusion imaging, *PloS One*, *5*(12), e15,710.

Fischl, B. (2012), FreeSurfer, *Neuroimage*, *62*(2), 774–781.

Flegal, J. M., and J. Hughes (2012), *mcmcse: Monte Carlo Standard Errors for MCMC R package version 1.0-1.*

Flegal, J. M., M. Haran, and G. L. Jones (2008), Markov chain Monte Carlo: Can we trust the third significant figure?, *Statistical Science*, pp. 250–260.

Friston, K., P. Jezzard, and R. Turner (1994a), Analysis of functional MRI time-series, *Human Brain Mapping*, *1*(2), 153–171.

Friston, K. J., A. P. Holmes, K. J. Worsley, J.-P. Poline, C. D. Frith, and R. S. Frackowiak (1994b), Statistical parametric maps in functional imaging: a general linear approach, *Human Brain Mapping*, *2*(4), 189–210.

Friston, K. J., A. P. Holmes, J. Poline, P. Grasby, S. Williams, R. S. Frackowiak, and R. Turner (1995), Analysis of fMRI time-series revisited, *Neuroimage*, *2*(1), 45–53.

Friston, K. J., D. E. Glaser, R. N. Henson, S. Kiebel, C. Phillips, and J. Ashburner (2002), Classical and Bayesian inference in neuroimaging: applications, *Neuroimage*, *16*(2), 484–512.

Fujiwara, N., K. Sakatani, Y. Katayama, Y. Murata, T. Hoshino, C. Fukaya, and T. Yamamoto (2004), Evoked-cerebral blood oxygenation changes in false-negative activations in BOLD contrast functional MRI of patients with brain tumors, *Neuroimage*, *21*(4), 1464–1471.

Gelfand, A., and P. Vounatsou (2003), Proper multivariate conditional autoregressive models for spatial data analysis, *Biostatistics*, *4*(1), 11–15.

Gelfand, A. E., L. Zhu, and B. P. Carlin (2001), On the change of support problem for spatio-temporal data, *Biostatistics*, *2*(1), 31–45.

Gelfand, A. E., H.-J. Kim, C. Sirmans, and S. Banerjee (2003), Spatial modeling with spatially varying coefficient processes, *Journal of the American Statistical Association*, *98*(462), 387–396.

Gelman, A., and X. L. Meng (1998), Simulating normalizing constants: From importance sampling to bridge sampling to path sampling, *Statistical Science*, *13*(2), 163–185.

Gelman, A., and D. B. Rubin (1992), Inference from iterative simulation using multiple sequences, *Statistical Science*, pp. 457–472.

Genovese, C. R., N. A. Lazar, and T. Nichols (2002), Thresholding of statistical maps in functional neuroimaging using the false discovery rate, *Neuroimage*, *15*(4), 870–878.

Gomori, J. M., R. I. Grossman, C. Yu-Ip, and T. Asakura (1987), NMR relaxation times of blood: dependence on field strength, oxidation state, and cell integrity, *Journal of Computer Assisted Tomography*, *11*(4), 684–690.

Gotway, C. A., and L. J. Young (2002), Combining incompatible spatial data, *Journal of the American Statistical Association*, *97*(458), 632–648.

Gross, W. L., and J. R. Binder (2014), Alternative thresholding methods for fMRI data optimized for surgical planning, *NeuroImage*, *84*, 554–561.

Håberg, A., K. A. Kvistad, G. Unsgård, and O. Haraldseth (2004), Preoperative blood oxygen level-dependent functional magnetic resonance imaging in patients with primary brain tumors: clinical application and outcome, *Neurosurgery*, *54*(4), 902–915.

Haller, S., and A. J. Bartsch (2009), Pitfalls in fMRI, *European Radiology*, *19*(11), 2689–2706.

Harms, M., J. Xu, E. Yacoub, D. Nolan, and D. Barch (2013), Impact of multiband EPI acquisition in a simple FMRI task paradigm and analysis, in *19th Annual Meeting of the Organization for Human Brain Mapping, Seattle, Washington, USA*, p. 3448.

Hastings, W. K. (1970), Monte Carlo sampling methods using Markov chains and their applications, *Biometrika*, *57*(1), 97–109.

Hickok, G., and D. Poeppel (2007), The cortical organization of speech processing, *Nat. Rev. Neuroscience*, *8*, 393–402.

Hirsch, J., et al. (2000), An integrated functional magnetic resonance imaging procedure for preoperative mapping of cortical areas associated with tactile, motor, language, and visual functions, *Neurosurgery*, *47*(3), 711–722.

Holmes, A., and K. Friston (1998), Generalisability, Random Effects & Population Inference, *Neuroimage*, *7*, S754.

Holmes, A. P., R. Blair, G. Watson, and I. Ford (1996), Nonparametric analysis of statistic images from functional mapping experiments, *Journal of Cerebral Blood Flow & Metabolism*, *16*(1), 7–22.

Hsu, Y. Y., C. N. Chang, S. M. Jung, K. E. Lim, J. C. Huang, S. Y. Fang, and H. L. Liu (2004), Blood oxygenation level-dependent MRI of cerebral gliomas during breath holding, *Journal of Magnetic Resonance Imaging*, *19*, 160–167.

Jenkinson, M., P. Bannister, M. Brady, and S. Smith (2002), Improved optimization for the robust and accurate linear registration and motion correction of brain images, *Neuroimage*, *17*(2), 825–841.

Jin, X., B. P. Carlin, and S. Banerjee (2005), Generalized hierarchical multivariate CAR models for areal data, *Biometrics*, *61*(4), 950–961.

Johnson, T. D., Z. Liu, A. J. Bartsch, and T. E. Nichols (2013), A Bayesian nonparametric Potts model with application to pre-surgical FMRI data, *Statistical Methods in Medical Research*, *22*(4), 364–381.

Kekhia, H., L. Rigolo, I. Norton, and A. J. Golby (2011), Special surgical considerations for functional brain mapping, *Neurosurgery Clinics of North America*, *22*(2), 111–132.

Kim, D.-S., T. Q. Duong, and S.-G. Kim (2000), High-resolution mapping of iso-orientation columns by fMRI, *Nature Neuroscience*, *3*(2), 164–169.

Kim, H., D. Sun, and R. K. Tsutakawa (2001), A bivariate Bayes method for improving the estimates of mortality rates with a twofold conditional autoregressive model, *Journal of the American Statistical Association*, *96*(456).

Krüger, G., and G. H. Glover (2001), Physiological noise in oxygenation-sensitive magnetic resonance imaging, *Magnetic Resonance in Medicine*, *46*(4), 631–637.

Kwong, K. K., et al. (1992), Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation., *Proceedings of the National Academy of Sciences*, *89*(12), 5675–5679.

Lindquist, M. A. (2008), The statistical analysis of fMRI data, *Statistical Science*, *23*(4), 439–464.

Liu, Z., V. J. Berrocal, A. J. Bartsch, and T. D. Johnson (2016), Pre-surgical fMRI data analysis using a spatially adaptive conditionally autoregressive model, *Bayesian Analysis*, *11*(2), 599–625.

Mardia, K. (1988), Multi-dimensional multivariate Gaussian Markov random fields with application to image processing, *Journal of Multivariate Analysis*, *24*(2), 265–284.

Müller, P., G. Parmigiani, and K. Rice (2007), FDR and Bayesian multiple comparisons rules, in *Bayesian Statistics 8*, edited by J. M. Bernardo, M. J. Bayarri, J. O. Berger, A. P. Dawid, D. Heckerman, A. F. M. Smith, and M. West, pp. 349–370, Oxford Sci. Publ., Oxford Univ. Press, Oxford.

Murray, I., Z. Ghahramani, and D. MacKay (2006), MCMC for doubly-intractable distributions, in *Proceedings of the 22nd Annual Conference on Uncertainty in Artificial Intelligence (UAI)*, edited by R. Dechter and T. S. Richardson, pp. 359–366, Cambridge: Association for Uncertainty in Artificial Intelligence Press.

Nichols, T., and S. Hayasaka (2003), Controlling the familywise error rate in functional neuroimaging: a comparative review, *Statistical Methods in Medical Research*, *12*(5), 419–446.

Nossek, E., et al. (2013), Intraoperative seizures during awake craniotomy: incidence and consequences: analysis of 477 patients., *Neurosurgery*, *73*(1), 135–140.

Ogawa, S., T. Lee, A. Kay, and D. Tank (1990), Brain magnetic resonance imaging with contrast dependent on blood oxygenation, *Proceedings of the National Academy of Sciences*, *87*(24), 9868–9872.

Ogawa, S., D. Tank, R. Menon, J. Ellermann, S. Kim, H. Merkle, and K. Ugurbil (1992), Intrinsic signal changes accompanying sensory stimulation: functional brain mapping with magnetic resonance imaging, *Proceedings of the National Academy of Sciences*, *89*(13), 5951–5955.

Ojemann, J. G., J. W. Miller, and D. L. Silbergeld (1996), Preserved function in brain invaded by tumor, *Neurosurgery*, *39*(2), 253–88.

Parker, J. A., R. V. Kenyon, and D. E. Troxel (1983), Comparison of interpolating methods for image resampling, *Medical Imaging, IEEE Transactions on*, *2*(1), 31–39.

Penny, W. D., A. Holmes, and K. Friston (2003), Random effects analysis, *Human Brain Function*, *2*, 843–850.

Penny, W. D., K. J. Friston, J. T. Ashburner, S. J. Kiebel, and T. E. Nichols (2011), *Statistical parametric mapping: the analysis of functional brain images*, Academic Press.

Potts, R. B. (1952), Some generalized order-disorder transformations, *Proc Camb Philos Soc*, *48*, 106–109.

Reich, B. J., and J. S. Hodges (2008), Modeling longitudinal spatial periodontal data: A spatially adaptive model with tools for specifying priors and checking fit, *Biometrics*, *64*(3), 790–799.

Reuter, M., and B. Fischl (2011), Avoiding asymmetry-induced bias in longitudinal image processing, *Neuroimage*, *57*(1), 19–21.

Reuter, M., H. D. Rosas, and B. Fischl (2010), Highly accurate inverse consistent registration: a robust approach, *Neuroimage*, *53*(4), 1181–1196.

Reuter, M., N. J. Schmansky, H. D. Rosas, and B. Fischl (2012), Within-subject template estimation for unbiased longitudinal image analysis, *Neuroimage*, *61*(4), 1402–1418.

Revett, K. (2011), An Introduction to Magnetic Resonance Imaging: From Image Acquisition to Clinical Diagnosis, in *Innovations in Intelligent Image Analysis*, pp. 127–161, Springer.

Riffel, P., and H. J. Michaely (2013), Improved echo planar diffusion-weighted imaging of the head and neck using syngo ZOOMit, *Clin Oncol*, *2*, 68–71.

Rue, H., and L. Held (2005), *Gaussian Markov random fields: theory and applications*, CRC Press.

Shaby, B. A. (2014), The open-faced sandwich adjustment for MCMC using estimating functions, *Journal of Computational and Graphical Statistics*, *23*(3), 853–876.

Siero, J., N. F. Ramsey, H. Hoogduin, D. Klomp, P. R. Luijten, and N. Petridou (2013), BOLD specificity and dynamics evaluated in humans at 7 T: comparing gradient-echo and spin-echo hemodynamic responses, *PloS One*, *8*(1), e54,560.

Smith, S. M. (2002), Fast robust automated brain extraction, *Human Brain Mapping*, *17*(3), 143–155.

Smith, S. M., et al. (2004), Advances in functional and structural MR image analysis and implementation as FSL, *Neuroimage*, *23*, S208–S219.

Springer, C., C. Patlak, I. Palyka, and W. Huang (1999), Principles of susceptibility constrast-based functional MRI: The sign of the functional MRI response, *Functional MRI*, pp. 91–102.

Stippich, C. (2007a), *Clinical functional MRI: presurgical functional neuroimaging*, Springer.

Stippich, C. (2007b), Presurgical functional magnetic resonance imaging (fMRI), *Clinical Neuroradiology*, *17*(2), 69–87.

Strother, S. C. (2006), Evaluating fMRI preprocessing pipelines, *Engineering in Medicine and Biology Magazine, IEEE*, *25*(2), 27–41.

Sunaert, S. (2006), Presurgical planning for tumor resectioning, *Journal of Magnetic Resonance Imaging*, *23*(6), 887–905.

Tharin, S., and A. Golby (2007), Functional brain mapping and its applications to neurosurgery, *Neurosurgery*, *60*(4), 185–202.

Tieleman, A., K. Deblaere, D. Van Roost, O. Van Damme, and E. Achten (2009), Preoperative fMRI in tumour surgery, *European Radiology*, *19*(10), 2523–2534.

Ulmer, J. L., L. Hacein-Bey, V. P. Mathews, W. M. Mueller, E. A. DeYoe, R. W. Prost, G. A. Meyer, H. G. Krouwer, and K. M. Schmainda (2004), Lesion-induced pseudo-dominance at functional magnetic resonance imaging: implications for preoperative assessments, with discussion, *Neurosurgery*, *55*, 569–591.

Voyvodic, J. T., J. R. Petrella, and A. H. Friedman (2009), fMRI activation mapping as a percentage of local excitation: consistent presurgical motor maps without threshold adjustment, *Journal of Magnetic Resonance Imaging*, *29*(4), 751–759.

Wall, M. M. (2004), A close look at the spatial structure implied by the CAR and SAR models, *Journal of Statistical Planning and Inference*, *121*(2), 311–324.

Welvaert, M., J. Durnez, B. Moerkerke, G. Verdoolaege, and Y. Rosseel (2011), neuRosim: an R package for generating fMRI data, *Journal of Statistical Software*, *44*(10), 1–18.

Whittle, P. (1954), On stationary processes in the plane, *Biometrika*, pp. 434–449.

Wikle, C. K., and L. M. Berliner (2005), Combining information across spatial scales, *Technometrics*, *47*(1).

Woolrich, M. W., T. E. Behrens, C. F. Beckmann, M. Jenkinson, and S. M. Smith (2004), Multilevel linear modelling for FMRI group analysis using Bayesian inference, *Neuroimage*, *21*(4), 1732–1747.

Woolrich, M. W., T. E. J. Behrens, C. F. Beckmann, and S. M. Smith (2005), Mixture models with adaptive spatial regularization for segmentation with an application to fMRI data, *Medical Imaging, IEEE Transactions on*, *24*(1), 1–11.

Woolrich, M. W., S. Jbabdi, B. Patenaude, M. Chappell, S. Makni, T. Behrens, C. Beckmann, M. Jenkinson, and S. M. Smith (2009), Bayesian analysis of neuroimaging data in FSL, *Neuroimage*, *45*(1), S173–S186.

Worsley, K. J., S. Marrett, P. Neelin, A. C. Vandal, K. J. Friston, and A. C. Evans (1996), A unified statistical approach for determining significant signals in images of cerebral activation, *Human Brain Mapping*, *4*(1), 58–73.

Yoo, S.-S., C. R. Guttmann, and L. P. Panych (2001), Multiresolution data acquisition and detection in functional MRI, *Neuroimage*, *14*(6), 1476–1485.

Yoo, S. S., I. F. Talos, A. J. Golby, P. M. Black, and L. P. Panych (2004), Evaluating requirements for spatial resolution of fMRI for neurosurgical planning, *Human Brain Mapping*, *21*(1), 34–43.

Yue, Y., J. M. Loh, and M. A. Lindquist (2010), Adaptive spatial smoothing of fMRI images, *Statistics and Its Interface*, *3*, 3–13.